A/c

Recd & issued [illegible]

[illegible]

7/8/75

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd

EDUCATIONAL RESEARCH

Volume 16 Number 1
November 1973

Contents

Subject options in a comprehensive school *J. Price* 3

Subject choice in the sixth form: some comments arising from a recent study *J. Selkirk* 8

Open or closed sixth: some consequences of structural change in comprehensive schools
G. R. Neave 12

Three myths in educational research *J. Wilson* 17

Construction of a comprehensive reading programme from a combination of published reading schemes
M. T. Kendall 20

The pre-reading vocabulary of children leaving nursery school *H. Quigley* 28

Some determinants of teacher morale in Australia *G. M. Coverdale* 34

Authoritarian personality traits among students *J. Dunham* 40

Commitment to teaching of postgraduate and college-trained students
D. Hellawell and A. Smithers 46

Individuals and systems: aspects of educational issues in drug dependence *D. A. Lane* 52

The three-minute reasoning test: a re-evaluation *J. Hartley, F. W. Hogarth and R. L. Mills* 58

The maturation of spatial ability in geography *A. J. Heamon* 63

Book reviews 67

32

Diary N 905(3)
21/8/75
Bureau Edn'l Psy Research.

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT. OF W. BENGAL
LIBRARY
CALCUTTA-19

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 16 Number 1
November 1973

Bureau of Edn & Psy. Research
S. C. E. R. T.
Acc. J 823

Subject Options in a Comprehensive School

John Price, Hounsdown School, Hampshire

Summary

This paper describes a practical attempt to resolve conflicting subject options of fourth- and fifth-formers in comprehensive schools. The system was devised to give the greatest satisfaction of individual needs for 224 pupils, who, in addition to 16 compulsory periods, had a choice of six subjects to fill a further 24 periods each week. In all, 52 subject options were offered at three levels: GCE 'O' level, CSE and non-examination courses. By the use of punched cards, the conflict between different subject groups was resolved to give 89 per cent satisfaction of pupils' choice. Implications for a banded system, for combinations of popular subject choices and for further changes in the system, are discussed.

The object of the system

Believing with the Swiss psychologist Paul Tournier that 'to live is to choose' our main purpose in the system we have developed at Hounsdown for the compilation of option groups has been to treat all 224 boys and girls in our eight-form entry third year as individuals; thus making it possible for each pupil during his fourth and fifth years to follow a personal timetable suited to his own particular needs, interests and, most important of all, varying levels of ability in different subjects.

It is fairly common practice in many schools to arrange bands of general ability in an attempt to meet the needs of the individual to a reasonable degree. Experience has shown this to be a tolerably successful system, but within each band are always found individuals who are far from ideally placed for particular subjects. Our researches at Hounsdown have shown most convincingly that individual differences are such that they are not best served by a system of banding. Of our 224 pupils only 18 (eight per cent) have been recommended independently by heads of departments to follow a GCE 'O' level course in every one of their chosen subjects, whilst 117 (52 per cent) have been recommended to follow at least one GCE 'O' level course. Even before these figures became known we have been careful not to generalize, but to speak of 'John Smith—individual' rather than 'GCE and CSE types' or even worse, 'non-exam types'. At the lowest ability end only three pupils (1·3 per cent) have been independently advised by departmental heads to take no external examinations whatsoever. See Table 1 for the complete breakdown of statistics.

We have welcomed the raising of the school leaving age because for the very first time we are able to fully integrate the ROSLA-ites into the system without the disturbance of early leaving. In fact they will at no time have been separated from their fellow pupils and we do not anticipate having to segregate them in the future. Like most schools we have seen fit in the past to cater separately for 'the leavers'. Now we and they will be spared that. We leave it to others to conjecture who the ROSLA-ites would have been and to draw dividing lines separating them from the rest of school society.

Basically, each pupil during his fourth and fifth years spends 16 periods a week on compulsory subjects plus six chosen subjects allocated four periods each, making up a 40 period week. Whilst we speak loosely of a pupil's choice it should be emphasized that this is a guided choice. Professional careers advisers, parents and form teachers have all been brought into the consideration of the individual pupil's requests, and heads of departments, although unable to turn pupils away from their departments—but for minor exceptions, do have the final say concerning the level at which the subject may be studied.

At this point it may be worthwhile to remind the reader that the usual method of compiling option groups is to put subjects into groups in the hope that they have been placed in such a manner that the average pupil (be he average GCE or average CSE) in selecting one of each ends up with a 'liberal' or 'rounded' timetable. Often a great deal of intuitive thought backed by experience goes into the grouping of subjects. Unfortunately the individual's unique combination of needs have to be discounted and even if he gets a fairly satisfactory combination of subjects he rarely gets a satisfactory combination of levels.

The system explained

Subjects are offered at three levels and for the purpose of this article it is a convenient generalization to say that level 1 represents GCE 'O' level, level 2 CSE and level 3 non-externally examined courses. However this should not be taken as a strict classification. Heads of departments, having been clearly instructed that the responsibility for catering for the full ability range was theirs, were invited to write separate entries for the school prospectus describing or 'selling' their courses. These entries were restricted in length to 40 words. Three entries typifying the different levels are given:

Course Reference	*Course*
GE1	GEOGRAPHY. A systematic course covering regional geography, physical geography and world regions to suit the GCE syllabus; yet with the direct approach of field studies, youth hostel visits and mapping experience that makes it enjoyable and challenging as well as academically useful.
LI2	LITERATURE. This is a CSE Mode 3 course specially designed for Hounsdown students. It will consist of studies of personal relationships and of the challenges of 20th century life as spotlighted in the literature—novels, plays and poems of today.
DL3	DESIGN FOR LIVING. It is hoped that this course, covering many interesting topics will be of very wide appeal. The course will cover do-it-yourself jobs, consumer choice and discretion, planning and finding a home, feeding the family, dress and design.

Fifty-two courses in all were finally listed in the prospectus, 18 at level 1, 23 at level 2 and 11 at level 3. No prior conditions were made such as: 'If you opt for this you must also opt for that' or 'If you opt for this you must not opt for that' or 'You must opt for at least one science' or 'You must not take more than two crafts'. Thus all pupils could freely select any six from the 52 offered. (Students of mathematics will realize that 52 courses taken six at a time, i.e. $^{52}C_6$, gives over 25 million possible combinations!) The degree of responsibility and awareness displayed by the vast majority of pupils was agreeably high. A few needed particularly careful guidance. These tended to be the least mature rather than the least able. Furthermore it was obvious, and thought to be a good thing, that parents were having a great deal of say in what their youngsters were opting for.

Bound into the prospectus was an Initial Inquiry Sheet on which each pupil listed the references of the courses he wanted in order of preference and against which he had to obtain departmental heads' signatures. Parents' and

Table 1: Numbers and percentages of pupils, from a total of 224, selecting courses from the three levels offered

Number of courses selected	LEVEL 1 *GCE 'O' level courses* No. of pupils	 % of pupils	LEVEL 2 *CSE courses* No. of pupils	 % of pupils	LEVEL 3 *Non-exam. courses* No. of pupils	 % of pupils
1	38	17	29	13	31	14
2	17	8	22	10	26	12
3	11	5	33	15	19	8
4	11	5	37	17	5	2
5	22	10	43	19	8	4
6	18	8	39	17	3	1
Totals	117	52	203	90	92	41

form teachers' signatures showed approval of the final combination of six subjects in each individual's case.

A specially printed punch-card bearing the 52 possible courses was then prepared for each pupil. Each card simply bore the pupil's name and form and his six courses which were punched out. The cost of the cards was about £5 and a platform punch (a lifetime's investment) about double that figure. The 224 cards were then given a two-way analysis and the statistics obtained were set out so as to form *Conflict Matrix A*. Restriction on space does not allow the complete reproduction of the matrix here but a paradigmatic sample is shown in Table 2. In this sample the bi-letter codings represent: biology, chemistry, computer appreciation, fabric crafts, French, music and statistics. In the full matrix each axis bears the 52 courses under consideration.

The matrix is self explanatory, showing, for example, in the principal diagonal 28 pupils opting to study fabric crafts and 85 opting to study French. In the body of the matrix we see for example that there are 12 pupils wanting to take GCE chemistry and CSE French. The data from conflict matrix A were handled in two stages, *Stage 1* was to consider only the figures in the principal diagonal. We decided on the basis of these figures, the number of classes that could be 'afforded' for each subject bearing in mind the practical limitations imposed by staffing and accommodation. CSE French for example qualified for three classes, GCE and CSE music had to be combined into a single class and statistics could no longer be offered as only three pupils had opted for it. We ended up with a final total of 67 course options which required arranging in the six option groups. *Stage 2* concerned the main body of the matrix with its 1,326 entries or 'conflicts' ($^{52}C_2=1,326$) these were extracted and ranked thus:

Table 2: Paradigmatic sample of two-way analyses of options

Principal diagonal

	BI2	CH1	CQ2	FC3	FR2	MU1	MU2	SI1
BI2	54							
CH1	0	34						
CQ2	1	1	23					
FC3	5	1	4	28				
FR2	28	12	7	5	85			
MU1	0	2	0	0	4	8		
MU2	1	0	0	0	4	0	4	
SI1	1	1	0	0	3	0	0	3

Conflict Matrix A

1.	FR2	HI2	38
2.	HI2	PH2	33
3.	HI2	RU2	29
4.	BI2	FR2	28
5.	BI2	HI2	28
•	•	•	•
•	•	•	•
•	•	•	•
•	•	•	•
50.	BI2	TI2	13

Theoretically of course it was possible to go to the 1,326th place. We went to the 50th place (in fact this was further than was strictly necessary) as the vast majority of the 1,326 conflicts were insignificantly small or zero. A three-way punched card analysis was then made on these 'high conflict' combinations and the statistics obtained were set out so as to form *Conflict Matrix B*. Again a paradigmatic sample is shown in Table 3. In the full matrix the vertical axis carried the 50 two-way combinations and the horizontal axis the 52 courses listed in the prospectus.

Conflict Matrix B was then examined and the most frequently occurring three-way conflicts extracted and ranked as follows:

1.	BI2	FR2	HI2	14
2.	FR2	GE2	HI2	14
3.	BI1	CH1	FR1	13
4.	CH1	GE1	PH1	13
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
27.	ME2	PH2	TD2	10

The final compilation of the six option groups was facilitated by a simple mechanical device consisting of a board with six columns of cup hooks (representing the six option groups)

Table 3: Paradigmatic sample of three-way analyses of options

	BI2	CH1	CQ2	FC3	FR2
28 BI2—HI2	X	0	2	2	14
28 BI2—FR2	X	0	0	2	X
29 HI2—RU2	6	0	5	5	9
33 HI2—PH2	9	1	2	0	13
38 FR2—HI2	14	0	4	3	X

Conflict Matrix B

awaiting receipt of 67 marked discs representing the 67 course options that had been decided upon. The discs were then alloted to the option columns by a strictly statistical method. The first to be placed were BI2 FR2 HI2. Because of the high conflict these were 'regarded as 'incompatible' and placed in different option groups. The discs were worked through systematically: firstly the three-way conflicts, then the two-way conflicts and finally the remaining single discs were alloted to an option group. The rule observed the whole time being 'great conflict means incompatibility and different option groups, little or zero conflict means compatibility and the same option group'. Conflict Matrix A yielded the information for most of the placings.

When the placings were completed two questions had to be asked. Can these be staffed? Can these be accommodated? Surprisingly only two changes had to be made for these reasons. For example, one option group demanded three geography teachers at a time when we could only produce two. So one of the courses had to be changed to the next best position. One change had to be made involving accommodation. By chance or design (I suspect the latter) the option groups turned out to contain roughly the same number of courses in each.

Success of the system

Having compiled the option groups we were anxious to know to what extent the pupils would be satisfied. One hundred per cent satisfaction was known to be impossible as that would have meant all 224 pupils getting precisely the six courses for which they had opted. The results in fact were:

95 obtained 6 out of 6
113 obtained 5 out of 6
15 obtained 4 out of 6
1 obtained 3 out of 6

224

These results are best expressed as percentage satisfaction.

$$\begin{aligned} 95 \times 6 &= 570 \\ 113 \times 5 &= 565 \\ 15 \times 4 &= 60 \\ 1 \times 3 &= 3 \\ & \overline{1{,}198} \end{aligned}$$

$$224 \times 6 = 1{,}344$$

Percentage satisfaction

$$= \frac{1198}{1344} \times 100 = 89\%$$

This remarkably high percentage satisfaction figure was very encouraging but we knew it was still subject to two factors:

(a) the position would be worsened if some pupils had to be moved out of oversized groups—this proved to be not the case.

(b) the position would be improved if disappointed pupils obtained a near substitute—and this in fact was the case.

Some observations

Whilst it is readily admitted that our system of treating every pupil as an individual and of taking the whole year intake as a single band falls short of perfection, it is difficult to see what advantages a multi-banded system has, unless it be less bother to organize. By putting the responsibility for backward and difficult pupils squarely into the hands of departmental heads we have avoided the problem of having to find and appoint that ephemeral person 'the teacher who is an expert with Newsom kids and ROSLA-ites'! Furthermore under our system such pupils never form into permanent groupings. Just before our system was launched a pupil asked me the question, 'Will I be able to do geography and history, because in some schools you have to do one or the other?'. The pupil is the daughter of a schoolmaster and I unhesitatingly answered 'Yes'. In fact the two-way analysis showed this particular combination to be one of the most popular (32 per cent of pupils) and it should give those schools which force a choice between the two cause for reflection. The question is sometimes asked whether our system allows for a change of mind by the pupil. Here too the answer is 'Yes', and I believe we can tolerate as much as ten per cent change without too much disruption. In fact requests for change have amounted to 5·6 per cent. Next year heads of departments of popular subjects may well offer courses at more than three levels whilst those of a more restricted appeal may decide on fewer than three levels. It will be apparent to the reader that the projected common examination at 16+ will suit our purposes even more admirably than the present GCE/CSE system. Furthermore if the percentile range were to be 35 to 100 rather than the proposed 40 to 100 I suspect that in our particular school every single pupil would be working towards at least one public examination. In other words it seems to me that every child might be at least a 35 'percentiler' in something.

Subject Choice in the Sixth Form

some Comments Arising from a Recent Study

J. Selkirk, Gosforth County Grammar School, Northumberland

Summary

The article draws attention to severe faults in the design of the Oxford (1960) enquiry into Arts/Science specialization. Research since then (recently reviewed by Lewis, 1972) has not been unanimous in refuting the hypothesis that the most able pupils are one-sided in their subject interests.

Evidence is offered on specialization and related issues from recent surveys (1970, '71 and '72) of the great majority of upper sixth formers in Northumberland. Findings included the following.

1. Subject choice at Advanced level is largely felt to be freely made rather than enforced.
2. Attitudes to wider courses are adverse rather than favourable.
3. High ability pupils are relatively concentrated in specialist courses, particularly the Sciences.
4. Many specialist pupils have a very wide range of achievement at Ordinary level.
5. The more able pupils make earlier subject choice decisions than the weaker: Science specialists show earliest decision times.
6. Subject choice is seen as easy rather than difficult by three-fifths of the sample pupils, and by three-quarters of Science specialists.
7. Science subjects are perceived, by either sex, as more closely interrelated than Arts subjects.
8. Mathematics is unusual in that it consistently attracts higher rates of dislike than other subjects, though four-fifths of those taking it would still select it if a choice had to be made again.
9. Definitions of 'specialization' need reconsideration.

Introduction

There can be little doubt that a central issue in the field of sixth-form subject choice is that of whether the most able pupils are markedly one-sided in their subject interests. This hypothesis was one of the themes of the Crowther report of 1960, and was also that which the study carried out in 1958 by the Department of Education of Oxford University is widely thought to have discredited (Oxford University Department of Education, 1960). In the intervening years, several workers have sought to clarify the many controversial features of sixth-form specialization, yet the recent critical review by Lewis (1972) in this journal, of six major researches, reveals that the crucial questions remain largely unresolved. It therefore seems appropriate to offer here some further evidence on two counts. First, as the Oxford investigation was that which most sharply called in question the concept of 'subject-mindedness', attention will be drawn to some apparently unreported features of that study. Second, surveys in Northumberland County Grammar schools, and in direct grant schools in the area, have provided data bearing on specialization for over 1,000 upper sixth-form pupils in each of the

---

Acknowledgements

The writer wishes to acknowledge the support of the University of Newcastle upon Tyne, and the co-operation of the Director of Education for Northumberland, together with heads of schools which co-operated in the enquiry mentioned here.

He also records his indebtedness to Dr. N. M. Newman of the Newcastle University Computing Laboratory, for programming and computing, and to Mr P. J. Wallis of the School of Education for many valuable suggestions and comments.

years 1970 and 1971, and brief mention will be made of what seem to be the more noteworthy findings.

The Oxford study

This was carried out by many members of the Department of Education, under the direction of A. D. C. Peterson, and funded by a grant from the Gulbenkian foundation. Whilst Lewis drew attention in his review of research to some limitations of the Oxford study, there were in fact several others which to this writer's knowledge have not been pointed out. The following are suggested to be the most serious.

1. Distinction must first be made between the actual 'Report', which was in four sections, discussing among other things historical origins of specialization and some European practices in it, and what Peterson termed the 'Enquiry'. The latter involved the completion of a questionnaire by a national (stratified random) sample of almost 3,000 sixth-formers, but was not reported upon separately: rather, references to it were interwoven with other themes. Yet though the 'Enquiry' must have provided, by virtue of its scale, a great deal of factual information, the report contains, at most, eight references to survey results, not all drawn from the national sample and not all concerning major issues. This curiously slight use of obtained data must strongly weigh against any claim of the investigation that the assertions made in it were founded on detailed experimental evidence.

2. The best-known finding of the enquiry was that, under a 'free choice' situation involving three Advanced level subjects, 40 per cent of pupils would choose mixed Arts and Sciences, 37 per cent would specialize in Arts and 23 per cent would do so in Science. The present writer has suggested elsewhere (Selkirk, 1972) that the wording of the question used—which began by asserting that pupils' choices were 'largely determined by outside requirements such as University entrance—seemed dangerously likely to pre-judge the issue of whether subject choices were 'free' or 'forced'. When Northumberland pupils were asked how they regarded their choices in this respect, roughly three-quarters—in each of two large samples—recorded choice as freely made. Whilst this result seems to support the suggestion that the original question wording was unsound, an alternative explanation is that the considerable increase in the numbers of subjects offered at Advanced level by schools today, compared with that of a decade ago, has indeed made possible something approaching genuinely 'free' choice. On the other hand, if this alternative is correct, it should be seen in the context that over half the three-subject pupils in the samples are still following rigidly specialist courses.

3. As Lewis pointed out, the Oxford enquiry had no means of classifying pupils on some measure of ability. This omission must be regarded as a severe fault in design, and from this fault derived several others. Perhaps the worst of these was that, although pupils were asked to name the *three* subjects they would select in a free choice situation, there must surely have been some *two*-subject pupils in the sample: indeed, a properly drawn stratified random sample would have been certain to contain some of these. It is idle now to speculate on what this proportion was, and the original records are believed not to have been kept. But it is certainly irrelevant, in an enquiry into specialization, to enquire which three subjects would be taken by those who aspired only to two, and it is likely also to be misleading. For such pupils would, it will be accepted, be the weaker ones, and therefore the more likely to choose mixed subject courses, thus artificially depressing the specialist proportions.

4. Taking further the theme of ability, Peterson remarked that at Ordinary level in his sample, 'the average pupil offered seven subjects and passed in six. Nearly 90 per cent of the sample passed in at least one Science and one Arts subject, well over half in at least two subjects on each "side", and a substantial number in at least three of each.' These findings were the basis for the assertion that no Arts/Science polarization exists at this stage. But as it was conceded that Ordinary level passes are evidence of no more than 'a fairly high degree of general intelligence', the recorded spread of passes—with no indication of high or low grades—is clearly a quite inadequate basis for the assertion made. It may be added that in work still in hand on the Northumberland data—which used mean Ordinary level grade as an ability/achievement index—the Ordinary level results of a sample of eighty pupils holding eight or more passes showed strong evidence of Arts/Science bias, which was highly correlated with type of eventual Advanced level specialization.

In 1969 Elton remarked that 'the fact that (Peterson's findings) have not been challenged by any subsequent study, in spite of the fact that they went against the most powerful tradition of the English secondary school, should convince an unbiased observer that they are hard to challenge effectively'. But it now seems more realistic to say that there were such severe shortcomings in that enquiry that the question of challenge may not seriously arise. Moreover, there is steadily accumulating evidence, from carefully designed researches, which does not all support Peterson's beliefs. It is true that he was at pains to point out that he believed specialization 'by sides' to be more harmful to the average than to the very able pupil. But caveats such as this were often overlooked, and, the provenance of the report and the unquestioned liberality of its author being what they were, it was widely felt to have been established that, in some undefined way, the mixing of Arts and Science subjects was 'the better thing' than specialization in either. Yet the growing evidence from recent studies may, in fact, have stemmed from the Oxford enquiry, for Peterson wrote that 'subject-mindedness . . . is a myth: if it is not, let us have some evidence for its existence'. Perhaps it was this challenge, together with other factors such as the evidence of the Dainton report, which led to an increasing interest in determinants of subject choice. And it may be that current and future researches in this field will lead in due course to that coherent and consistent model of subject choice which we so sadly lack at present.

The Northumberland surveys

In 1970 the writer was able, through the support of the Newcastle University Department (now School) of Education to undertake an enquiry covering many aspects of 'A' level subject choice. The target group was the entire upper sixth form population of Northumberland, together with that of several direct grant schools in the area. The questionnaire used consisted of some forty items, many involving subjective judgements in their responses, and, to guard against serious error in variation arising from this source, the entire survey was repeated in 1971. Nineteen schools took part, providing an acceptable balance of the sexes, and yielding response rates each year of over 80 per cent. The information obtained was extensive enough to need computer assistance in its analysis, and here the Social Science Research Council supported the enquiry with an award for data-processing costs.

Analysis to date provides some interesting results relating to Arts/Science specialization[1]. Of particular interest was the finding that high-ability pupils were relatively concentrated in specialist courses, especially in Science. Mixed subject combinations contained some high-ability pupils but were characterized rather by their high proportion of pupils with very modest achievements. It was also found that almost two-thirds of those pupils holding nine or more 'O' level passes were following specialist courses. Such pupils are very likely to have both Arts and Science subjects well represented, though it is often alleged that Advanced level specialization is the outcome of too many options closed before Ordinary level.

Attitude to wider sixth-form courses was found to be adverse rather than favourable, though this is a difficult issue to explore, as question forms tend to be worded in terms of a mixture of examinable and non-examinable subjects. With regard to the stage in school at which subject choices are made, there was strong evidence that the more able pupils took such decisions earlier than the weaker. The more able were also found to be very much more likely than the weaker to record 'their own enquiries and general reading' as the major determinant, among several sources of advice, of subject-choice decisions. Science specialists showed earlier decision times than other pupils: this agrees with Hudson's comment (1966) that 'there is some evidence that young scientists commit themselves to science at a younger age than arts specialists commit themselves to the Arts'.

Of especial interest were results relating to whether Advanced level subject choice was found to be easy or difficult. Some three-fifths of the samples recorded it as easy, boys finding it more so than girls: almost three-quarters of Science specialists found it not to be a problem. Approaching this theme from a rather different angle, pupils were asked whether they made their choices by deliberately selecting related subjects, or simply on the basis of the interest and appeal of each one. Here emerged one of a great many sex differences found, namely, that boys much more commonly than girls recorded choices made by related subjects rather than individually. Subsequent analysis revealed that this result is partly due to the male/female pattern long established in Science/Arts specialization, and further suggested that Arts subjects are not perceived by pupils as being inter-related to anything like the extent that Science subjects are. The sex difference is not the dominant effect here, for Science girls very largely recorded choice by related subjects and Arts boys did not. It is relevant here to point to one of the less tractable—and least discussed—features of any research on specialization, namely, the problem of defining it. Thus, the combination of English, history and French would generally be accepted as an instance of Arts specialization. Yet there is a very genuine sense in which the 'mixed' group of history, mathematics and economics is much more specialized, particularly if the history is modern (or social) and the mathematics involves the statistics option. We may need to reconsider currently accepted classifications on a basis of 'inter-relatedness', if we are not to miss some major clues to the unravelling of the mysteries attending specialization.

---

[1] Further information to be included in a higher degree thesis, can be obtained from the author.

A further enquiry in 1972 (School of Education, Newcastle) examined attitudinal variables and ability in all Northumberland pupils offering mathematics as one of their Advanced level subjects. Among many interesting results was that involving satisfaction with the choices made on entry to the sixth form. Five terms later, four-fifths of the sample recorded that they would still choose mathematics if they had their choice to make again, and this is especially significant in that mathematics had been previously shown to evoke less favourable attitudes than other subjects. Further studies are needed here, for it is often suggested that sixth form pupils are frequently discontented with their subject choices.

Conclusion

Lewis' review achieved the difficult task of briefly summarizing and relating several major studies. It contained also the reasoned argument that further research should be designed to study the determinants of subject choice by pupils *in the same school*, on the hypothesis that these determinants much outweigh inter-school differences 'of organization and ethos'. Another suggestion likely to be fruitful is that of Hudson, who urged the need for further study of the interests and perceptions of young children, and pointed out that there exists definite evidence that 'marked bents towards Science, for example, exist long before the specialist teacher gets to work'. These two suggestions emphasize the need for the study within schools of both psychological and sociological factors if further progress in our knowledge of subject choice is to be made.

References

CENTRAL ADVISORY COUNCIL FOR EDUCATION (1959). *15 to 18.* London: HMSO (The Crowther Report).

ELTON, L. R. B. (1969). 'The sixth form of the future—the headmasters' view', *Universities Quarterly*, Summer, 326.

HUDSON, L. (1966). *Contrary Imaginations*, 113. London: Methuen.

HUDSON, L. (1968). *Frames of Mind*, 61. London: Methuen.

LEWIS, D. G. (1972). 'Subject choice in the sixth form: a critical review of research', *Educ. Res.* **15**, 1, 21-7.

OXFORD UNIVERSITY DEPARTMENT OF EDUCATION (1960). *Arts and Science Sides in the Sixth Form.* A report to the Gulbenkian Foundation, Oxford.

PETERSON, A. D. C. (1960). 'The myth of subject-mindedness', *Universities Quarterly*, **14**, 3, 223.

SCHOOL OF EDUCATION, NEWCASTLE (1972). Report on an enquiry into attitude and performance among pupils taking Advanced level Mathematics. Unpublished.

SELKIRK, J. (1972). 'Subject choice at GCE Advanced Level', *Universities Quarterly*, Winter, 74-89.

Open or Closed Sixth

some Consequences of Structural Change in Comprehensive Schools

Guy R. Neave, Centre for Educational Sociology, University of Edinburgh

Summary

Based on a survey of 969 university entrants from 163 comprehensive schools in England and Wales, this article[1] examines the effect of the 'open' sixth form upon the educational background of university students. It finds that the 'open' sixth has considerable influence upon subject preference at 'A' level, upon the number of 'O' levels passed in the sixth form and upon the ability to change subjects at this late stage in a student's school career. The findings suggest that the 'open' sixth form is a far more flexible institution than its 'closed' counterpart, at least as far as university students are concerned. Finally, the article advances the idea that structural change in the sixth form is likely to be far more conducive to flexibility of subject choice than manipulation of the examination structure.

Introduction

During the past few years, the sixth form has been under pressure in three particular ways. Firstly, structural change; secondly, pressure for a revision of its curriculum; and thirdly, the outcome of the first two: the possible changes in the relationship with higher education brought about by structural or curricular alterations. These are not all the pressures at work by any manner of means. Nevertheless, they are the most important since they affect the nature of the links between secondary and higher education.

Structural alteration in the sixth form is best represented in the changed definition as to what constitutes a sixth former. The Schools Council (1966) accepts as sixth formers all those remaining in school after their 16th birthday. In addition to this widely accepted definition, recent research into the sixth form of comprehensive schools has drawn a more precise line between the types of sixth that are to be found in this sort of school. In their survey of comprehensive schools in Britain, Benn and Simon (1970) distinguished between 'open' and 'closed' sixth forms. Entry to 'closed' sixth forms is conditional upon a student gaining a stipulated number of 'O' levels or their equivalents. The 'restricted' or 'closed' sixth of the Benn-Simon enquiry had no students working for 'O' levels, whilst the 'open' sixth in addition to 'A' level students also accepted others whose sole educational goal was 'O' level. The change in structures involved in moving from a 'closed' to an 'open' sixth form is especially important in comprehensive schools. By 1970, 71 per cent of the schools survey by Benn and Simon had 'open' pattern sixths.

Structural changes of this kind are not without implication for curriculum reform since the setting up of the 'open' sixths extends the broad ability found in the age ranges 11 to 16 into the sixth form itself. To date, however, most of the proposals for revising the sixth-form curriculum —whether in the shape of Intermediate Level, Qualifying or Further Level—have concentrated on the needs of the university-bound student (Peterson, 1968). Curriculum reform at the post-16 level is still largely working on the assumption of a homogeneous ability range in the sixth form, though evidence on the evolution of that part of the school is largely to the contrary.

In view of the changes in the structure of the sixth form taking place in comprehensive schools it is important to see what effect, if any, structural alteration has upon the type of subject combinations passed at 'A' level by university entrants from both types of sixth form—'open' and

---

[1] The research report on which this article is based will be published next year by Routledge and Kegan Paul under the title *How They Fared: The Impact of the Comprehensive School upon the University.*

'closed'. For it is here at the point where sixth form and higher education interlock through examination structures that structural change could have far-reaching consequences. There are already some indications that comprehensive schools' sixth formers take different types of 'A' level combinations from those usually found in grammar schools. Morris' (1969) work taking sixth formers in Birmingham and Herefordshire schools suggested that more students from all-in schools sat a mixture of Arts and Science subjects at 'A' level. The question still remains, however, to what extent these patterns show up in the 'A' level combinations of university entrants from comprehensive schools and to what extent they are influenced by the structure of the sixth form.

Research Method

The enquiry was carried out in two stages. In the first a list of schools was compiled from the DES list of comprehensive schools existing in January 1967. Local Education Authorities were then asked to provide information on those students who entered university in October 1968 from those schools under their jurisdiction that figured on the DES list. In the case of large LEAs—London and Bristol for example—the schools were approached directly, otherwise data were furnished by the LEAs on the basis of their Grants List. The schools attended by university entrants were then located in the data files and tapes used in the original Benn-Simon investigation. The decision to study university entrants of October 1968 was governed by the fact that the Benn-Simon survey was conducted in July of the same year. Thus, information on the structure of the schools attended by individual students was accurate only for the class that graduated to university the following October. Of a possible 391 comprehensive schools in England and Wales with sixth forms, information on their university entrants was available for 163—that is 41·6 per cent of all comprehensives with sixth forms in existence in July 1968.

The second stage involved a one in two sample of the names given by the LEAs, selected by means of a random number indicator. Approximately 83 per cent replied (i.e., 969 questionnaires were valid and usable) and it was upon these data that the findings are based.

Data gathered by questionnaire were then collated to see what was known about schools attended by individual students: for example, whether the admission to the sixth form was dependent on a given number of 'O' levels prior to entry. This information, gathered in stage one of the enquiry, was then added to that given by each individual student, thus enabling us to control for such effects as admission procedures to the sixth form dependent upon the type of 'A' level combination passed, the number of 'O' levels passed whilst in the sixth form, etc.

Results

Of the 969 students, 347 or 35·8 per cent entered university from 'open' sixth forms. Just under three quarters (71 per cent) of 'open sixth formers' attended schools that had been comprehensive prior to, but no later than, September 1961.

The date when the comprehensive school received its first 'comprehensive intake' of students is an important factor. For only those students who passed the whole of their secondary school careers in a school designated 'comprehensive' can be said to be the first 'comprehensive generation' to reach university in the context of this survey.

The first area to be investigated was whether the 'open' sixth allowed more students to pass 'O' level after entry to the sixth form than the restricted or 'closed' sixth form. One enquiry conducted on behalf of the Schools' Council (1971) revealed that comprehensive school sixth formers tended to pass more 'O' levels in the sixth form itself than students from any other type of school. No indication, was given, however, as to whether this pattern was due to the admissions policy of the sixth, or whether more students who failed on their first attempt were resitting 'O' levels at that stage in their careers. Nor, naturally, was there any indication as to how this might affect potential university entrants. In the present survey, students were asked to show at what stage in their school careers they passed individual 'O' level subjects. Fifty-two per cent passed at least one 'O' level whilst in the sixth form. At this stage, the admissions policy to the sixth appeared to have no perceptible effect. 53·3 per cent of 'open' sixth formers passed at least one 'O' level in the sixth and 51·3 per cent of 'closed' sixth formers. The results are reported in Table 1.

It is not known whether these 'O' levels were taken for the first time in the sixth, or whether

Table 1: Number of 'O' levels passed in the sixth form by type of sixth

No. of 'O' Levels	*'Open' Sixth*	%	*'Closed' Sixth*	%
1	74	40·0	195	61·1
2	79	42·7	87	27·3
3	15	8·1	26	8·2
4	11	5·9	6	1·9
5+	6	3·2	5	1·6
N =	185	100·00	319	100·00

$\chi^2 = 24{\cdot}55$ $df = 4$ p less than 0·001 $C = {\cdot}214$

they were repeats and therefore allowance cannot be made for any difference in the figures in Table 1. But it would appear that the results are generally in line with those of the Schools' Council sixth form survey—with this difference—that the 'open' sixth is just as important for future university students who need to gain the necessary basic qualifications to be eligible for a place in higher education as it is for the terminal student. That a significantly higher proportion of 'open' sixth formers going to university passes two 'O' levels in the sixth form means, in effect, that the 'open' sixth appears to bring 'O' and 'A' level considerably closer together in time. Any future revision of the sixth form curriculum will have to bear this in mind. In other words, any future proposal for a two-tier examination structure for the sixth on the lines of the Q and F pattern with one examination for school leavers and another for university entrance, both to be taken in the sixth form, will demand the abolition of 'O' level. Or, if not its abolition, then its confinement to the lower school.

To what extent 'open' sixth formers pass more 'O' levels whilst actually in the sixth because of flexibility or of any important curricular changes following on the structural redefinition of the sixth form is not known, though there is some evidence in the data, not discussed here, that suggests it is due to a larger than average number of 'A' level subjects offered. Nevertheless, there are certain indications that structural redefinition does have most specific effects upon the 'A' level pattern of students in the 'open' type sixth form. The 'A' level subjects passed by the 969 students were combined into four categories—Arts, Science, Social Science and Mixed. The latter group comprised at least one science subject combined with either Arts or Social Science 'A' levels. Students mixing their 'A' levels across the Arts/Science boundary have been the centre of considerable interest since the publication of the Dainton Report (Great Britain, 1968) which pointed out the possibility that this type of student could well double over the period from 1966 to 1973.

Accepting that this is so, and a recent DES survey (1967) appears to bear it out, the precise conditions in the sixth form that create this 'A' level hybrid are by no means clear. However, since the rationale of the 'closed' sixth form is closely associated with the subject requirements of higher education, our hypothesis was that the 'open' sixth form would permit a much freer permutation of subjects even at 'A' level, because it has to deal with a far more varied range of ability than the closed sixth. The complementary hypothesis was that the 'closed' sixth would be more likely to concentrate its students on the normally accepted straight arts or sciences. In order to test this hypothesis out, students passing specified subject combinations at 'A' level were grouped according to whether entry to the sixth form in their school was 'open' or 'closed'. The results are shown in Table 2.

The main areas of difference are that 'open' sixth formers tend to move away from straight science subjects. And the move into mixed 'A' levels is quite clear, being almost the double of the pattern shown by 'closed' sixth formers. Thus, insofar as the presence of students with mixed 'A' levels is an indication of flexibility, the 'open' sixth form would appear to be more flexible than the 'closed' sixth. As for the complementary hypothesis, 'closed' sixth formers tend to con-

Table 2: 'A' level subject combinations passed by university entrants from comprehensive schools with 'open' and 'closed' sixth forms

	'Open' Sixth	%	*'Closed' Sixth*	%
Arts	73	21·0	145	23·3
Science	162	46·7	315	50·6
Social Science	47	13·5	102	16·4
Mixed	65	18·7	59	9·5
Unknown			1	·2
N =	347		622	

$\chi^2 = 17{\cdot}72$ $df = 3$ p less than ·001 $C = {\cdot}10$

centrate to a greater degree on Arts and Science combinations—74 per cent passing in these two groups—than students in 'open' sixth forms of whom 68 per cent passed the same two combinations.

A further test of flexibility is, of course, the ability of the school to allow students to change their courses in the sixth form. Accordingly, students were asked 'At any time in your sixth form 'A' level course did you change from the subject(s) you had originally intended to study?'. Only a minority did so—12 per cent or 124 students. Even so, far more 'open' sixth formers said they had changed course than students in 'closed' sixths—8·0 per cent (50 students) of the latter compared to 18·4 per cent (64 students) from the former ($\chi^2 = 23{\cdot}23$, *df* 1, p less than 0·001, $C = {\cdot}14$).

Discussion

The survey on which this article is based was limited to university entrants—terminal sixth formers were not considered. This, of course, limits the degree to which one can generalize about the effect of sixth form structures on the latter group. But it is probably true to say that if the structure of the sixth form has pronounced effects upon university entrants, it will not have any less effect upon students not continuing their education full-time beyond 18. It would appear from these results that the 'open' sixth form is bringing about some significant changes as regards the type of 'A' level combination passed by high ability students and in the possibility of changing course in mid-stream as well as acting as a significant 'fail-safe' device for students to make up or to repeat 'O' levels at that point in their school careers. Peterson (1960) suggested that a high proportion of sixth formers would, in effect, prefer to combine Arts with Science 'A' levels, but pointed out at the same time the lack of any institution or body at national level capable of bringing about the necessary flexibility in the curriculum of the sixth form to realize this hidden student demand. More recently, proposals presented by the Schools' Council have sought to introduce more flexibility at sixth form level by altering the pattern of sixth form courses, partly through wider combinations, partly through a redefinition in the content of specific subject areas.

But it appears that a certain measure of flexibility in choice is attained through structural change at sixth form level, even within the confines of the present 'A' level syllabus. Flexibility in options and subject choice at sixth form could well turn out to depend on two major variables with a third as possible intermediary. The first of these is the structure of the sixth form—whether 'open' or 'closed'—the second, the content and permutation of courses available. The third intermediary variable might possibly turn on the question of size of sixth form and its rate of expansion. Obviously, the larger the sixth form, the greater the number of subjects that can be supported without violating the judicious use of teacher resources. And the quicker the expansion, the more easily this situation can be reached. In this respect therefore, the 'open' sixth would appear to have far greater potential for rapid expansion than the 'closed' sixth, especially with the raising of the school leaving age.

On the other hand, where the sixth form structure is 'closed' and its curriculum associated in the main with Advanced level and entry to higher education, flexibility of choice might best be introduced through change in the examination structure and in the subject content of specific areas of study. In contrast, where the structure of the sixth is expanded to include terminal students whose educational goals might only be one or two 'O' levels, the necessity of having to cater for a wide range of ability might well result in greater flexibility of choice for students of high ability. If this is the case, then the 'open' sixth would appear in a far better position to generate its own internal flexibility rather than relying on outside manipulation through the examination system.

Since the history of curriculum reform at sixth form level is littered with spectacular rejections, any acceptance is likely to be an unconscionable time a-coming. In the meantime, schools wishing to impart added flexibility to their sixth form curriculum could well do so by changing over to the 'open' sixth.

References

BENN, C. and SIMON B. (1970). *Half Way There.* Maidenhead: McGraw Hill, 187–90.

DEPARTMENT OF EDUCATION AND SCIENCE (1967). 'Sixth form survey', *Trends in Education*, No. 7.

GREAT BRITAIN, COUNCIL FOR SCIENTIFIC POLICY (1968). *Enquiry into the Flow of Candidates in Science and Technology into Higher Education* (Dainton Report). London: HMSO.

MORRIS, R. N. (1969). *Sixth Form and College Entrance*. London: Routledge & Kegan Paul, 158.

PETERSON, A. D. C. (1960). *Arts and Science Sides in the Sixth Form*. A Report to the Gulbenkian Foundation, Oxford.

PETERSON, A. D. C. (1968). 'Sixth form future', *Education*, November 29th, 1968, 652.

SCHOOLS COUNCIL (1966). *Working Paper No. 5. Sixth Form Curriculum and Examinations*. London: HMSO.

SCHOOLS COUNCIL (1971). *Sixth Form Pupils and their Teachers, Vol. 1*. London: Books for Schools.

Three Myths in Educational Research

John Wilson, Department of Educational Studies, Oxford University

Summary

There are certain assumptions, often unconscious, which dominate a lot of educational research. Not everyone shares them, but they are widespread enough to need explication. I think that few people except the demonstrably lunatic would share them when explicated; but they are usually too vague and subterranean to be easily pinned down. Hence I call them 'myths'.

1. The questionnaire or 'hard facts' myth

Asking people why they do things, even in a formal questionnaire, is less grotesquely behaviouralistic than measuring movements or stimuli; but it is tarred with the same brush. An example: several researchers have published detailed statistics on the percentages of parents, teachers and others who approve or disapprove of 'R.E.' and 'moral education'. Any sensible person will at once wonder *what it is* that they are asked to approve or disapprove under these titles: might not different respondents mean different things by 'R.E.'? This turns out to be the case, as my colleagues and I found in some (unpublished) research at Oxford in 1970-71. 'Unstructured interviews in depth' (a grand name for just *talking* to people) showed quite clearly that different people had very different, and often diametrically opposed, concepts or pictures sparked off by 'R.E.', 'moral education' and other such terms. Hence the respondents were, in effect, being asked totally different questions. A lurid sub-example: respondents A and B both said they approved of 'sex education' in schools. Cross-questioned, A turned out to have a picture of 'sex education' in which the idea was essentially to repress pupils, 'keep them pure', etc., which A thought was a good idea: B turned out to have a sort of *Oz* or D. H. Lawrence-like picture of it, whereby kids were encouraged to have all sorts of sexual experience ('After all, that's education, isn't it?'), which B thought was a good idea.

So, obviously, we have to find out the respondents' concepts of any terms we use in a questionnaire. This is often extremely difficult, and involves—something missing from almost all psychological research—producing an adequate taxonomy of concepts or pictures to begin with: and this in turn involves a lot of hard, not to say philosophical, thinking on our part. Things get worse when we ask respondents, for instance, why they chose to do this or that. We want to ensure that their answers are honest (difficult but not impossible), and also that they are giving their *real* reasons, not those that they have deceived themselves into accepting as their reasons (impossible by questionnaire methods only).

The only answer is to accept research methods which are not structured in *this* way. The myth is that this must result in 'soft' facts, mere 'subjective' impressions, etc.; but there are, of course, extremely rigid standards of procedure in this kind of research—as clinical psychologists, anthropologists, and other workers who try to find out what is really going on in a person's head (so to speak) very well know. We might then have some chance of designing questionnaires which tell us something interesting about people.

2. The 'replicability' myth

Borrowed from the physical sciences, with the disastrous inappropriateness typical of such borrowings, is the notion that any genuine or reputable piece of research must be 'replicable'. In one sense this notion is harmless. Any explanation is of the general form 'X behaved in way Y because of P and Q'; and any explainer thereby commits himself to the opinion that, *other things being equal*, P and Q will on any occasion make X behave in way Y. This sort of 'replicability' is part of the concept of an explanation.

Naturally we try to keep 'other things' as equal as possible (some call this 'controlling the variables'); but even in the physical sciences a total replication of any situation is virtually impossible. Yet much work has been done without this: consider astronomy, zoology and many other branches of science. What we do, of course, is to develop a refined picture of the field—planets, animals, or whatever it may be—in which we *guess* ('hypothesise') what the relevant factors are: colour and size are discounted, mass and skeletal structure seen to be relevant. We can get a long way in determining 'generative mechanisms' (see Harré and Secord, 1972, on this and the subject generally) without anything like an *experimentum crucis* or total replicability. To put it another way: we could not get a proper idea of what 'other things being equal' *meant* in particular contexts—of which 'other things' were relevant and which irrelevant—if we did not *first* fight our way towards a fairly high level of understanding.

A fortiori this is true of the human sciences. The notion of replicability in anthropology, for instance, is plainly absurd: yet equally plainly what the anthropologist learns about one tribe may help our understanding of tribes in general—just as (a point one might well have remembered right at the start) it is only by knowing *some* people that, in our everyday life, we come to have more general knowledge of people as a whole: without any question of replicated experiments. The same is true of the historian and the clinical psychiatrist.

The temptation, I think, is to insist on some particular concept of replicability ('Will it work with any children? With any teacher?'), and then to tailor our research interests to fit this model. Not surprisingly, there are few exciting results. What we should rather do is to ask the difficult but important questions, and tailor our research methods (including *some* concept of replicability, if you like) to fit these. To try to make psychology like physics is not the only defence against bias, prejudice, or lack of rigour.

3. The 'value-free' myth

If the idea was simply that the researcher should not incorporate his own particular values, preferences, prejudices or what you will into research, this would not be a myth. Nor, I suppose, would it be a myth if researchers simply *accepted* the 'values' from politicians, teachers, fashion, or some other source: I should be inclined to call this more of a disgrace ('The Führer knows that it's valuable for all kids to grow up hating Jews: just get to work and tell us how to do this'). The myth is that one can do serious research without oneself tackling the question, just by restricting oneself to 'the facts'.

Consider research into reading. We want 'good readers' or 'competent readers'. But what is to *count as* a 'good reader'? If you like grand language, this is a 'question of value': and now, the myth goes, either we must duck this question or else we shall only be using our own 'value-judgements'. Nonsense: what we have to do is to *tackle* the question: to look at the possible concepts involved. For instance, are we interested only in producing or finding out who *can* read or (also) who actually *does* read? Are we to mean by 'a good reader' someone who only reads when under extreme pressure to do so? Again, are we interested only in correct phonetic interpretation of written signs (only the signs used in the English language?), or also in whether he *understands* what he reads? If the latter, how *much* understanding? Understanding of what sort of subject-matter? All these questions arise, and more: to try and do 'research into reading' without considering them properly is just crazy—we wouldn't know *what* research we were doing.

The point is that it isn't entirely arbitrary ('a value-judgement') how these questions are answered. We do, in a semi-conscious way, have *some* idea of what we want to mean by 'a good reader', 'a morally educated person', 'being able to think like a competent scientist', 'breaking down class barriers', 'a sense of failure' and so forth. But this needs to be made conscious and precise, by comparing possible concepts and interpretations. Having done this, it will usually be obvious what we are after: but even if there are possible choices to be made, at least we shall be clear what these are. We can say, for example, 'If by "morally educated" you mean PQR, then we'll do one type of research: if you mean XYZ, we'll do quite another type'.

Getting the 'values' *clear* and precisely expressed is not only a preliminary to making 'value-judgements'—though it is at least that. Rather, it is the proper way of determining what our values and desires actually are. I do not deny that there can be differences of opinion; but we do not have either to make some arbitrary choice, or to accept some arbitrary ruling. In research, it is usually not so much a matter of deciding whether we want pupils to be 'mature', 'responsible', 'well-educated', and so on: it is a matter of getting clear about what we mean by these terms. Nobody, for instance, wants children to be stupid: but when we come to consider what is *meant by* 'intelligent' we have a difficult task—yet, obviously, an essential one if we are to assess or promote 'intelligence' in any normal sense.

What I have written above is (by design) brief, and necessarily sounds a bit dictatorial; though interested readers may pursue some relevant points elsewhere (Wilson, 1972). Any value it may have will accrue only if researchers are enabled to recognize remnants of these myths in their thinking, and thereby to expel them more easily. This needs courage as well as clarity. I should guess that the apparent unreality of much educational research, as seen by teachers and other practical workers, stems from these defects. Of course it is easy for philosophers to stand on the touch-line and criticize; but then that is what philosophers are for, and not much is gained by making them drink hemlock.

References

HARRÉ R. and SECORD, P. (1972). *The Explanation of Social Behaviour*. Oxford: Blackwell.

WILSON, J. (1972). *Philosophy and Educational Research*. Slough: NFER.

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Editor: John McLeish, Ph.D.
Educational Psychology Department
University of Alberta, Edmonton

Contents of September 1973 issue
Restricted, Extended Codes and Socioeconomic Status ☐ Ethnic Segregation in Toronto ☐ Open-Book Examinations ☐ Personalities of Male and Female Student Teachers ☐ Conceptual Tempo ☐ Adult Education ☐ Cognitive Abilities in Relation to Environment ☐ Reviews

Contents of December 1973 issue
Use of Drugs by University Students ☐ Intelligence Tests in Alberta ☐ Disadvantaged Canadian Indian Children ☐ Orientation of Soviet Teachers ☐ Family Size and Intelligence ☐ Analysis of Musical Form ☐ Reviews

Subscription: $6·00 (£2·50) per annum
Back issues of volumes I-XVIII available at special rates

Construction of a Comprehensive Reading Programme

from a Combination of Published Reading Schemes

Mary T. Kendall, Ryefield Infant School, Hillingdon

Summary

The teaching of reading is complicated by the fact that no one commercially produced set of 'readers' is adequate in all respects for the development of reading skills. The article which follows provides guidelines by which a complete reading scheme can be developed from a combination of reading schemes; including consideration of such fundamentals as compatibility of vocabulary, degree of repetition, phonic criteria and the selection of supplementary reading material.

This article has been prepared as a result of the author's conviction that no single available reading scheme alone provides sufficient material in the depth or variety necessary to effect the acquisition of reading skills by the normal child. The objective of the study described below was to establish a reading system which would, as nearly as possible, provide a continuous progression of reading material suitable for the wide range of ability of children passing through an infant school. For obvious reasons, the reading material needs to be selected from that which is commercially available and marketed as purportedly complete reading schemes.

A basic 'look and say' vocabulary

The first task was to establish a series of books which would provide a basic 'look and say' vocabulary of about 100 words; giving at the same time sufficient repetition for easy assimilation by the normal child. A further requirement was to supplement this series of books with additional material of similar vocabulary for the less able child and to establish in the programme an order for the teaching of phonics which could be used as a readability guide for further reading material. The 37 reading schemes which were available are listed in Goodacre (1968). Each one of these schemes was studied to determine those which had the greatest degree of compatibility in their vocabulary. Books were compared one with another in a trial and error system with like words being noted. The three reading schemes with a most closely related vocabulary proved to be 'Janet and John' (Nisbet), 'Ladybird' (Wills & Hepworth) and 'McKee' (Nelson). The books from these three schemes were placed in an order such that the new vocabulary introduced by each new book was reduced by approximately 50 per cent.

Table 1 summarizes an analysis of the vocabulary lists which were prepared for each of the listed books and places the books in the order in which they should be read, showing the new sight ('look and say') and phonic vocabulary for each book and the number of revision words which, together, give the total vocabulary of that book within its own scheme. It should be appreciated that the stated number of revision words relates to the combined scheme which uses three sets of readers. Within individual schemes the number of revision words at each stage would, generally, be significantly less except, of course, in the case of the 'Janet and John's' revision vocabulary in the *Little Books*. Though the 'McKee Platform Readers' have no listed vocabulary they repeat

Bureau of Edul. & Psyl. Research
S. C. E. R. T.)
te
J 823
BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT OF W. BENGAL
LIBRARY
CALCUTTA-19

Table 1: A combined reading scheme [compiled from three reading schemes: 'Janet and John' (Nisbet), 'The McKee Readers' (Nelson), 'Ladybird' (Wills & Hepworth).]

Reader Sequence Number		*Scheme Vocabulary* New Words			
		Sight	Phonic	Revision Words	Total Words
	Stage I				
1	'Here We Go'	27	0	0	27
2	'Little Books' *1–5*	0	0	27	27
3	'Off to Play'	32	0	0	59
4	'Little Books' *6–10*	0	0	32	59
5	'McKee 1' *Tip*	10	0	10	69
6	'Ladybird' *1a and 1b*	6	0	10	75
7	'Ladybird' *2a and 2b*	15	0	12	90
	Stage II				
8	'Little Books' *11–15*	22	11	13	123
9	'Out and About'	0	0	33	123
10	'McKee 2' *Tip and Mitten*	14	3	13	140
11	'McKee Platform Readers' *A1–A6*	—	—	—	—
12	'Ladybird' *3a and 3b*	12	6	18	158
13	'Little Books' *16–32*	15	56	11	229
14	'I Went Walking'	0	0	71	229
	Stage III				
15	'McKee 3' *With Peter and Susan*	11	3	22	243
16	'McKee Platform Readers' *B1–B6*	—	—	—	—
17	'Ladybird' *4a and 4b*	7	7	27	257
18	'Through the Garden Gate' and 'Blue Books'	35	169	26	461
19	'Ladybird' *5a and 5b*	2	14	30	477
20	'McKee 4' *Up and Away*	18	51	78	546
21	'McKee Platform Readers' *C1–C6*	—	—	—	—
	Stage IV				
22	'Ladybird' *6a and 6b*	5	11	36	562
23	'Ladybird' *7a and 7b*	14	21	32	597
24	'McKee 5' *On We Go*	0	62	151	659
25	'McKee Platform Readers' *D1–D6*	—	—	—	—
26	'I Know a Story'	2	287	70	948

words used in the main scheme and are, therefore, included in Table 1. As they also have many phonic words, such words have been added to the relevant phonic lists.

Having established a 'look and say' base of 90 words, the rest of the 26 reading levels were divided into four stages and the vocabulary for each stage listed and assessed phonically. Words repeating phonic elements were extracted and listed in Table 1 as 'phonic' words if, at that point, the child following the phonic programme of the combined scheme could be expected to tackle it on a phonic basis.

Supplementary reading for the less able

The next task was to identify a series of books with similar vocabulary for supplementary reading by the less able child. To do this the vocabulary for each of the four stages of the combined scheme summarized in Table 1 was listed alphabetically, each word retaining as a reference the reader sequence number of the book in which it first appeared. Books from other reading schemes were then checked against this vocabulary list and given the number of that book, followed by a + sign, after which, should it be read, it would provide a minimal new vocabulary. Lists of such supplementary reading books appear in Table 2. Proper names were included in the number of new words listed. If proper names were 'told' instead of being taught then the new vocabulary would be reduced, e.g. *We Look and See* (Wheaton) at 2+ would introduce only two words—'oh' and 'funny'. The books in Table 2 were drawn from a further eight reading schemes. Where books belong to the same reading scheme and have the same + number then the earlier book would be read first. As these books are placed in an order related solely to vocabulary it should be appreciated that some will not be of the correct interest level for many of the children and would be excluded from their reading programmes.

Phonics

The next step was to find an order for the teaching of phonics which would relate to the books being read, and to use this order to assess phonically further reading books with which to extend and enrich the child's vocabulary. Thus phonic criteria could be used to assess books where no obvious similarity existed in the vocabularies of the various books. For example, the scheme 'One, Two, Three And Away' (Rupert Hart Davis) appears to be incompatible with the combined scheme vocabulary on a 'look and say' basis, but, after study, it became obvious that provided a child is first taught the phonics for Stage 2 as described below he would be able to read Books One, Two and Three of that scheme. To establish a programme for the teaching of phonics which would relate to the books which the child would be using, some simple rules had to be postulated. So, having become familiar with the vocabulary of the various stages, words were listed according to the following rules.

1. Two-, three- and four-letter words consisting only of consonants and short vowels.
2. Consonant blends and consonant digraphs.
3. Double vowels.
4. Final *e* for long vowel sounds.
5. Dipthongs.
6. Vowel modifications with *r*.

Table 2: Supplementary Reading Books

Reader Sequence Number	*Description*		*Relationship to Scheme Vocabulary*		
			New Words		Revision
	Book	Publisher	Sight	Phonic	Words
2+	'Happy Trio' *We Look and See*	Wheaton	8	—	5
2+	'Kathy and Mark'	Nisbet	12	—	17
4+	'Happy Trio' *We Work and Play*	Wheaton	11	—	11
4+	'Happy Trio' *We Come and Go*	Wheaton	8	—	13
4+	'Kathy and Mark' *2*	Nisbet	22	—	13
4+	'Through the Rainbow' *Red 1*	Schofield & Simms	18	—	10
	(including conversation words) *Red 2*	Schofield & Simms	6	—	23
	Red 3	Schofield & Simms	10	—	27
7+	*Kitty and Rover* 'Beacon'	Ginn	11	—	29
7+	'New Vanguard Readers' *Bk. 1*		10	8	28
7+	'Happy Venture' *Fluff and Nip*	Oliver & Boyd	9	11	24
7+	'Happy Venture' *Bk. 1*	Oliver & Boyd	17	20	25
7+	'Springboard Readers' *1*	Warne	8	1	12
	'Springboard Readers' *2*	Warne	6	2	18
7+	'The Ginger Books' *Introductory*	Educational Supply Association	9	4	12
	'The Ginger Books' *1*	Educational Supply Association	11	5	39
	'The Ginger Books' *2*	Educational Supply Association	19	3	8
	'The Ginger Books' *3*	Educational Supply Association	12	5	5
7+	'The Adventure of Betty & Colin' *Bk. 1*	Davis & Moughton	7	9	12
	'The Adventure of Betty & Colin' *1a, b, c & d*	Davis & Moughton	12	0	—
	'The Adventure of Betty & Colin' *Bk. 2*	Davis & Moughton	17	9	14
7+	'Ready to Read'	Methuen			
	1. *Early in the Morning*		7	2	9
	2. *Grandma comes to stay*		10	1	7
	3. *The Fire Engine*		10	1	7
	4. *Where is Timothy?*		7	2	6
	5. *Going to School*		20	1	3
	6. *Playtime*		8	2	6

7. Soft *c* and *g* before *e* and *i*.
8. Silent letters.

The vocabulary of the books in Table 1 was listed on a phonic basis, each word being given the number of the stage in which it first appeared. It was then possible to define the phonic teaching which would help at each stage. It was interesting to note that some phonic elements, e.g. *oi* and *aw*, did not appear at all. Obvious omissions of this kind were added to the stage containing similar sounds: at Stage 3, *oi* with *oy* and at Stage 4 *aw* with *au*. If a phonic element appeared at least three times it was defined as a teaching point for that level. Where several consonants blended with *l* or *r* appeared, all such combinations were listed to be taught at that stage though some blends appeared only once in the whole scheme. Having established the phonics teaching programme in this way, it was possible to use the phonic criteria at each stage as a readability guide to obtain further enriching reading material from 33 other schemes or sets of readers.

Where the words appeared outside the phonic criteria they were discounted if (1) they were repeated sufficiently or (2) had good picture or contextual clues.

The phonic criteria were not an exhaustive list.

Stage 1

By the end of Book Seven it is assumed that most children will have acquired a knowledge of the sounds of the alphabet and can synthesize two-, three- and four-letter words containing short vowels. It will be obvious that supplementary readers for Stage 1 need to relate more to 'look and say' vocabulary than to phonics, for the child must have time to acquire sounds and their synthesis, and books containing similar 'look and say' words and words exemplifying these phonic criteria are rare.

Stage 2

The phonic criteria are as follows:

sh th st ch qu
endings: *-ed -all -nd -ng -nk*
oo (soon) *oo* (wood) *ee*, *ea* (eat) *oa*
Consonant blends with: *r cr gr pr fr br dr*
l cl gl pl fl bl sl
sw sn sm sk sp tr
wh

Less than 50 books were found for this level. This was obviously inadequate as a library of books for the beginner reader who had already 'read' 14 reading primers. A study was made of the first books of the seven most popular reading schemes (as listed in Goodacre, 1967) so that each scheme provided as near as possible a vocabulary of 100 words.

The combination of 667 words, the total of the seven schemes, resulted in an effective vocabulary of 259 words. When these words were listed according to their popularity in the seven reading schemes some interesting facts emerged. Not one noun was common to all the schemes and only the following ten words were found in every scheme: *a, and, can, in, is, it, the, to, up, with.* More books containing early 'look and say' words in different contexts are necessary for the beginner reader to allow him to gain confidence in his growing skill.

Stage 3

At Stage 3 the addition of the following phonics greatly increased the number of suitable books.

Phonic criteria:
Mute *e* ending modifying *a, i, o, u.*
a, e, i, o, u, modified by *r, ir, ur, er, or, ar*
oi, oy, *ow* (how), *ow* (own) *ou* (out)
Soft *c*, *ce, ci, gh, kn, -ing, y* in merry.

It is now possible to extend and enrich the child's vocabulary. It was interesting to note that many books in school libraries with an interest level of five years appeared on this list, and that the 'Beacon Reading Scheme' (Ginn) which incorporates phonics in the early stages could only appear at this stage of the combined reading scheme with Book 1, Part 1 and the rest of the books at Beacon Book One stage.

Books One to Twelve of the 'Griffin Readers', the A, B and C 'Dragon Books', the 'Sea Hawk Books' One and Two and 'Cowboy Sam Books' One to Six, all published by E. J. Arnold, are examples of books which can now be read together with some of the 'Rescue Reading' books (Ginn).

Stage 4

At Stage 4 books are plentiful

Phonic criteria:
aw, au, soft *g* *y-ie* *gu*
silent letters *t, b, w, l, g*
ee (head) *ie* (pier) *ph* *-tion* *-sion*
spr, squ, thr, spl.

The phonic lists from which the criteria for

Table 3: Relationship between 'Breakthrough to Literacy' vocabulary content and the early books of various reading schemes

Book	*Publisher*	*Scheme Vocab.*	*'Breakthrough to Literacy' Vocab.*
'Here We Go'	Nisbet	95	59
McKee	Nelson	86	53
'Ladybird'	Wills & Hepworth	79	46
'Kathy and Mark'	Nisbet	110	55
'Through The Rainbow'	Schofield & Sims	105	60
"Time For Reading'	Ginn	117	64
'Ready To Read'	Methuen	109	52
'Betty and Colin'	Davis & Moughton	93	42
'Queensway'	Evans	111	53
'Warne Springboard Readers'	Warne	99	56
'Beacon'	Ginn	114	48
'Dominoes'	Oliver & Boyd	104	51
'Racing To Read'	E. J. Arnold	94	47
'Happy Venture'	Oliver & Boyd	106	48

supplementary readers were drawn showed that even using a combination of three popular reading schemes, practice for many phonic elements was not possible, nor was there evidence of repetition for 'phonic' words containing similar elements.

Teachers using controlled vocabulary reading schemes are left to remedy their phonic deficiencies very often with only phonic lists at the back of the book which contain words bearing very little relation to the book's vocabulary. Phonics learned by the child should assist in his reading and words learned in this process should relate to his reading books. Phonic schemes such as 'Gay Way' (Macmillan) and/or 'Royal Road Readers' (Chatto & Windus) will fit easily into the phonic stages outlined, with the addition in the case of Royal Road of further phonic elements. The S.R.A. 'Word Games' and the 'Word Study Kit' for use with the Bell & Howell 'Language Master' can be used in the order dictated by the school's phonic programme.

The reading programme described above was developed whilst the author was actively engaged in class teaching. Since the work was completed additional reading schemes have been published: 'Breakthrough to Literacy' (Longmans), 'Dominoes' (Oliver & Boyd) and 'Sparks' (Blackie). These more recently available reading schemes can be integrated in the same way as that outlined.

The first words a child learns to 'read' should be vitally interesting to him, and the teacher by her choice of reading materials can choose which words these should be. Whatever she uses, be it 'Breakthrough to Literacy', conventional reading schemes, or her own list and reading materials, she is imposing her order of words on the child, for without order and repetition of words and/or sounds, a 'look and say' vocabulary and/or phonic competence cannot be acquired by the majority of children.

On a 'look and say' basis a comparison of approximately 100 words of two combined reading schemes may be of interest. The first set of words is that of the combined reading scheme which is the subject of this article; the second set is that composed of words found in reading schemes allied to 'Breakthrough to Literacy', the latter scheme chosen because it gives both teacher and child the freedom to compose more interesting sentences and to manipulate the word order of the language. Table 3 compares the relationship of 'Breakthrough to Literacy' words used and the early vocabulary (approximately the first 100 words) in 14 reading schemes, and shows that almost 50 per cent of these first words are also 'Breakthrough to Literacy' words. A child having a thorough knowledge of the words on his *Sentence Maker* (in the 'Breakthrough to Literacy' kit) would have no problem in reading the early books of most reading schemes. It is unlikely however that a child conversant with the vocabulary on his Sentence Maker would want to read these early readers, yet whilst he is acquiring his first words he wants to 'read' published books.

The 'Breakthrough to Literacy' books, delightful because they lack the boring repetition of many reading schemes, were used only as supplementary readers because they use a very wide vocabulary.

The *Red* and *Yellow* 'Breakthrough to Literacy Books' contain totals of from 21 to 79 words which include a range of up to 35 words which are not included in the 'Breakthrough to Literacy' vocabulary.

Combined Reading Scheme (2)

The reading scheme chosen as a base for the second combined reading scheme was 'Through the Rainbow' (Schofield and Sims) because its early books contain many words in common with 'Breakthrough to Literacy'. The first two books of the following schemes, 'Early to Read', 'Racing to Read' (both by E. J. Arnold) and 'Dominoes' (Oliver & Boyd), were added.

A summary of the vocabulary for this scheme, designated Combined Scheme 2, showing the number of words common to the 'Breakthrough to Literacy' vocabulary and the new words for each book appears in Table 4.

A comparison of the vocabulary lists for each of the Combined Schemes is interesting. The early vocabulary of the Combined Scheme, Vocabulary List 1, relates to kittens and dogs running, jumping and playing; aeroplanes, squirrels, birds and balls going up and down; rides on horses and in aeroplanes or big red boats. Teachers using apparatus and kinaesthetic aids in the teaching of vocabulary are often reinforcing words which the child must reproduce in his so-called 'free writing' simply because he cannot produce any others.

Vocabulary List 2, being allied to 'Breakthrough to Literacy', gives the child the opportunity to compose his own sentences before he is able to write. The first words relate to mummy and daddy working in the house and garden; the baby and the dog playing with toys and the child involved in various activities in school. Such words are often required by the child in his free writing; and other words on the Sentence Maker plus personal words give the child the opportunity to extend his vocabulary in a way which is interesting to him. It is for the teacher to choose the kind of early words that she will teach.

Vocabulary list: Combined scheme 1

STAGE I

1. 'Here We Go'
2. 'Little Books' *1–5*
3. 'Off to Play'
4. 'Little Books' *6–10*
5. 'McKee' *Tip*
6. 'Ladybird' *1a and 1b*
7. 'Ladybird' *2a and 2b*

1. aeroplane
and
3. a
at
7. are
1. boats
3. basket
big
5. ball
1. can
come
dog
down
3. fast
father
fly
for
5. find
7. fish
fun
1. go
3. goodbye

1. here
horse
5. home
6. has
7. have
he
1. I
3. in
is
it
7. into
1. Janet
John
jump
6. Jane
1. kitten
little
look
3. let
like
6. likes
1. my

3. may
me
mother
5. no
not
1. one
3. on
play
5. Peter
7. Pat
1. ride
run
3. red
1. see
3. said
skip
slide
5. Susan
6. shop
7. says
some
sweets

1. the
three
two
3. thank
this
to
too
6. toy
5. Tip
6. top
tree
7. they
1. up
3. us
very
want
went
5. will
with
7. wants
water
we
3. you
7. yes

Table 4: Combined reading scheme 2

Reader Sequence Number	Description	New Vocabulary	Revision Words	Total new vocab. in own scheme	'Break through to Lit.' vocab. inc. in New Vocab.
	'Through the Rainbow' (Schofield & Sims)				
1	*Picture Book 1*	15	vocab. unlisted	15	12
2	*Picture Book 2*	7	vocab. unlisted	7	6
3	*Picture Book 3*	4	vocab. unlisted	4	2
4	*Red Book 1*	9	19	9	6
5	*Red Book 2*	9	3	9	7
6	*Red Book 3*	6	4	6	3
	'Early to Read' (E. J. Arnold)				
7	*Early to Read Book 2*	6	11	17	2
8	*Early to Read Book 3*	11	7	18	6
	'Racing to Read' (E. J. Arnold)				
9	*Racing to Read Book 1*	10	14	24	3
10	*Racing to Read Book 2*	3	9	12	2
	'Dominoes' (Oliver & Boyd)				
11	*Dominoes Book 1*	13	11	24	6
12	*Dominoes Book 2*	12	4	16	5
		105			60

Vocabulary list 2: Combined list 2

1. *Picture Book 1*
2. *Picture Book 2*
3. *Picture Book 3*
4. *Red Book 1*
5. *Red Book 2*
6. *Red Book 3*
7. *Early to Read Book 2*
8. *Early to Read Book 3*
9. *Racing to Read Book 1*
10. *Racing to Read Book 2*
11. *Dominoes Book 1*
12. *Dominoes Book 2*

*Denotes 'Breakthrough to Literacy' vocabulary and derivatives.

1.	*a
1.	*and
9.	*are
11.	*at
1.	*baby
1.	*bed
5.	*be
6.	*being
7.	blue
8.	ball
8.	*by
9.	*big
11.	bricks
11.	*books
12.	boxes
3.	*can
7.	*car
9.	chimney
12.	clay
1.	*daddy
1.	*dog
1.	dolls
2.	dinner
3.	*do
3.	does
4.	draw
7.	dress
9.	door
11.	dressing
12.	*day
12.	else
12.	every
1.	*for
8.	father
9.	flowers
10.	frock
1.	garden
5.	*go
9.	green
8.	*get
1.	*house
3.	here
4.	*he
4.	help
6.	*have
7.	*has
1.	*in
2.	*is
5.	*I
8.	*it
11.	look
1.	*mummy
4.	*my
4.	*me
2.	*making
5.	*make
8.	mother
9.	*of
2.	*playing
5.	*play
5.	put
6.	*party
10.	*pretty
12.	*paint
12.	pictures
7.	red
8.	*runs
9.	roof
12.	*read
4.	*she
8.	*sees
11.	*school
11.	*school
11.	sand
12.	sing
12.	story
1.	*the
1.	toys
4.	*there
5.	*to
5.	tea
6.	table
7.	train
8.	*they
8.	tree
8.	tent
11.	*today
12.	things
12.	things
6.	under
11.	*up
2.	*working
2.	*with
2.	*what
4.	*write
5.	*work
9.	window
9.	wall
10.	*we
11.	water
1.	*you
4.	your
6.	yourself

Reading schemes cited in the text

Title	*Publisher*
1. 'Beacon Reading Scheme'	Ginn
2. 'Breakthrough to Literacy'	Longmans
3. 'Cowboy Sam Books'	E. J. Arnold
4. 'Dominoes'	Oliver & Boyd
5. 'Dragon Books'	E. J. Arnold
6. 'Early to Read'	E. J. Arnold
7. 'Gay Way'	Macmillan
8. 'Griffin Readers'	E. J. Arnold
9. 'One, Two, Three and Away'	Rupert Hart Davis
10. 'Racing to Read'	E. J. Arnold
11. 'Rescue Reading'	Ginn
12. 'Royal Road Readers'	Chatto & Windus
13. 'Sea Hawk Books'	E. J. Arnold
14. 'Sparks'	Blackie

References

GOODACRE, E. (1967). *Reading in Infant Classes.* Slough: NFER.

GOODACRE, E. (1968). 'A list of published reading schemes for the primary school'. Privately published, now out of print, and superseded by

ROOT, B. (1972). 'Learning to Read: a catalogue of books for all ages and stages'. Centre for The Teaching of Reading, University of Reading.

Journal of Curriculum Studies **Collins**

In the November issue

Conventional Classrooms, "Open" Classrooms and the Technology of Teaching

Devising Evaluation Instruments for Technological Problem-solving

Survey of History Teaching in Selected Secondary Schools

Some Problems Raised by the Schools' Council Humanities Project

There are also reviews of some of the most recent books in the field of curriculum studies

It may be ordered direct from the Publishers.

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

Annual Subscription £2.00 net including postage

In the U.S.A. and Canada
Agathon Publications Services Inc.
150 Fifth Avenue
New York, N.Y. 10011

Annual Subscription $7.50 including postage

The Pre-reading Vocabulary of Children Leaving Nursery School

Helen Quigley, formerly Research Officer, NFER Pre-school Project

Summary

A group of children, in their last term of attendance at nursery schools, were tested on their comprehension of words used in the first books of well-known reading schemes. No child recognized all the words, and the vocabularies of individual children were seen to differ greatly. The implications of this finding for infant and nursery school teachers are discussed.

Introduction

The impetus for this small study came from the author's need to assess the linguistic attainments of nursery school children as part of the groundwork for instituting a language development scheme in the National Foundation for Educational Research in England and Wales' Pre-School Project.

One of the aims of this project was to promote reading readiness in pre-school children. Language development is usually considered to be one of the more vital aspects of this readiness—the child must have an adequate vocabulary and be a master of basic syntactic structures if he is to cope with reading. Both Menyuk in America, and Widlake in England have demonstrated that the basic syntactic structures underlying a mature grammar are present in the speech of pre-school children, whether they be from cultured or disadvantaged backgrounds (Menyuk, 1963, 1964; Widlake, 1971). However, Widlake has also shown that in the case of pre-school children from an Educational Priority Area this competence is coupled with a limited vocabulary. Compared with their middle class contemporaries these deprived children scored low on the English Picture Vocabulary Test (Brimer and Dunn, 1965), a standardized measure of listening, or receptive, vocabulary.

One of the bars to beginning reading seems to be that disadvantaged children are not familiar with the words they are being asked to decipher. It seemed desirable, therefore, to find out whether this supposition was justified and, if so, to find which words were involved. There would then be a strong case for including these words in any language development programme. A criterion reference test was therefore devised to sample the children's knowledge of words used at the beginning of their reading experiences. This test was given to 50 children in their last term of attendance at four Local Education Authority nursery schools in South-East England.

The children

The sole criterion for selection was that the children should be native English speakers in their last term of attendance at an LEA Nursery School. Their ages, therefore, ranged from 4 years 9 months to 5 years 2 months, with an average of 4 years 11 months.

Social class

It was not possible to obtain accurate information about the parental occupation in each case. However, nursery school teachers have very close contacts with the mothers of their children and from the teachers' knowledge of the background of the children the social class composition of the sample, by father's occupation, was judged to be:

Social Class	*I and II*	*III*	*IV and V*	*Unclassified*
Father's Occupation	Professional and Managerial	Upper Working Class	Semi-skilled and Unskilled Manual	Fatherless
n	2	29	13	6
%	4	58	26	12
Expected %	22	50	26	2

These figures show that the sample contains a larger percentage of children from broken homes and fewer middle-class children than in the national population (Registrar General, 1968). The difference is statistically significant (χ^2 test, $p<0{\cdot}01$) and would be expected from an LEA nursery school population where places are awarded on the basis of either social or medical need. This social class distribution is very similar to that obtained in the NFER Pre-School Project. However, none of the children in this study was involved in that Project.

Verbal ability

In order to assess the general level of receptive vocabulary in the sample (obviously relevant to this study) the children were tested with the English Picture Vocabulary Test 1 which is a British standardization of the American Peabody Picture Vocabulary Test. The task is a multiple choice one—the child pointing to that picture, from the four presented, which corresponds to the word spoken by the tester. As the stimulus words become increasingly complex the result gives a measure of the level of difficulty at which the child listens with understanding. Claims by Brimer and Dunn that the test is a measure of verbal ability are generally treated with caution but there is no doubt that it provides a rapid assessment of the child's understanding and the range of his vocabulary.

The group scored at an average level on this test with a mean score of 100 and standard deviation of 14. As the test is constructed to give a mean of 100 with standard deviation of 15 the group does not deviate significantly from the norm. As is usual with this test, the boys scored slightly higher than the girls, with an average score of 101·8, to the girls' average of 99·6. This difference is not statistically significant.

Scores of children aged below 5 years 0 months were obtained by extrapolation. This practice is not recommended by the constructors of the EPVT but it was felt to be better than testing the children on two separate tests.

We were thus reasonably confident that our group of children was representative, in both ability and background, of the Nursery School population.

The test

This was designed along the lines of the English Picture Vocabulary Test, and used some of the distractors from that instrument along with other pictures specially drawn for the purpose.

Content

The criterion for the study was the vocabulary content of first reading books. Twenty infant school headmistresses in the area were asked to name the reading scheme the majority of children in their school first encounter. From the replies it was clear that nearly all children entering infant school are faced with one of six reading schemes. These are listed at the end of the references.

As the test was solely intended to measure the vocabulary of beginning readers, we decided only to include the first 50 words from each of these books. As there was some overlap in the content of the books the potential total of 300 words reduced to the 134 different ones which formed the 'pool' of items for our test. These were then sub-divided into nouns (34 words), verbs (46 words), adjectives (12 words), articles (3 words), pronouns (14 words), adverbs and prepositions (13 words) and others (11 words—mainly conjunctions, exclamations, etc.). We decided that the format of the test meant that only present participles of verbs could be tested while most of the words categorized as 'others' did not prove suitable (e.g. 'Oh!', 'please', and 'ting'). Finally, a test was constructed covering 73 words—34 nouns, 19 verbs, 11 adjectives, 2 pronouns, 5 prepositions and adverbs and 2 'other' words. The plates representing these words and their distractors, were bound into a small booklet which was offered to the child in the same manner as the EPVT.

Testing

The testing was done by the author in the children's own schools. As the two tests were so similar there was a pause between them when the children engaged in activities of a different nature.

Testing children of this age is always an unreliable procedure—an attempt was made to assess the reliability of the instrument by giving the test on a second occasion, three weeks later, to 15 of the children. This gave a test-retest co-efficient of 0·92, which is satisfactorily high.

Results

The results were analysed in four separate categories:

(a) the total 'word knowledge' score of each child

(b) variations of this total with respect to sex and social class

(c) the familiarity of each word to the children

(d) variations in this familiarity between different parts of speech.

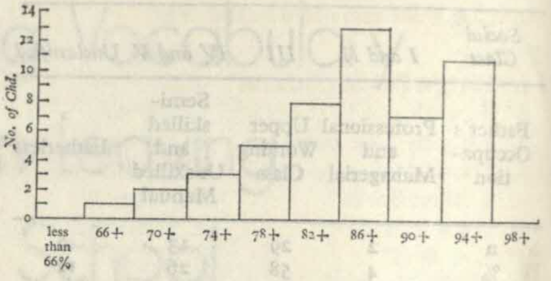
(*a*) *Total word-knowledge scores*

These were calculated by totalling the score each child obtained on the 79-word test and percentaging this score. The following distribution of total scores was obtained.

Percentage of words known	Number of children obtaining this score	Percentage of children obtaining this score
98–100	0	0
94–97	11	22
90–93	7	14
86–89	13	26
82–85	8	16
78–81	5	10
74–77	3	6
70–73	2	4
66–69	1	2
	50	100

The distribution of scores is obviously negatively skewed, as no child obtained a score lower than 66 per cent. The mean was 88 per cent. This is disconcertingly low in view of the vital nature of these words. This means that the average child leaving a good nursery school may be baffled by as many as 12 per cent of the words in his reading primer.

Within the group the distribution of scores also has a negative skew.



Of course, the sample is small, but this skew could indicate that some words have low discrimination values (i.e. children scoring high, may score erratically on these words). This would be a defect were one trying to standardize the test; however, in a test of this nature it merely highlights the individual differences in vocabulary that can occur at all score levels.

(*b*) *Variations in total score with sex and social class*

There was no significant sex difference in the results, mean scores for boys and girls being 88 per cent and 87 per cent respectively.

Social class analysis is hampered by the small sample size. However, the average score obtained by the nine children from social classes I, II and IIIa was 91 per cent while that of the 35 children from classes IIIb, IV and V (i.e. manual backgrounds) was 87 per cent.

(*c*) *The familiarity of each word to the children*

This was measured by counting the percentage of children who understood the word. The results are grouped by familiarity, and part of speech within each familiarity grouping.

(i) Words known by 100 per cent of the children

Nouns	*Verbs*	*Prepositions*	*Pronouns*	*Other*
Aeroplane	Bring	Under	Her	Miaow
Baby	Climb			Woof
Boat	Eat			
Boy	Feed			
Cat	Go			
Coat	Jump			
Dog	Play			
Fish	Read			
Penny	Ride			
Pig	Run			
Tree	Say			
	Sit			
	Slide			
	Walk			

(ii) Words known by 90–99 per cent of the children

Nouns	*Verbs*	*Prepositions*	*Adjectives*
Ball	Fall	On	Big
House			Fat
Ice Cream			Two
Kitten			Wet
Lorry			
Mother			
Pot			
Slide			
Tin			
Toy			
Water			

(iii) Words known by 76–89 per cent of the children

Nouns	*Verbs*	*Adverbs and Prepositions*	*Adjectives*
Fox	Meet	In	Bad
Hen	Skip	Top	Little
Horse			No
Hill			One
Sister			Pretty
School			Red

(iv) Words known by 50–75 per cent of the children

Nouns	*Verbs*	*Adverbs and Adjectives*	*Pronouns*
Can	Catch	Very	His
Home	Sea	Three	
Rat			

(v) Words known by less than 50 per cent of the children

Nouns
Cap
Fun
Mud

This indicates that 44 of the 73 words (i.e. 60 per cent) were not known by all the children and would therefore qualify for inclusion in a vocabulary expansion programme. Obviously, with over 90 per cent of the children knowing 63 per cent of the words the problem is not one of mass ignorance. It is rather than individual children have individual vocabularies which do not always include the words in common reading books. For instance, how many children regularly use the word 'mother'?

(d) Analysis of these by parts of speech

It is interesting to see whether the familiarity varies with different parts of speech. Infant teachers know that young children usually read the nouns and verbs in their books quickly—it is the 'little words' that cause trouble. Unfortunately, these same 'little words' proved difficult to sample within the format of our test. However, for the words we sampled, the percentage of words in each category known by all the children were:

Parts of speech	*Number of words in this category in the test*	*Number of words in this category known by all the children*	*Percentage of words in this category known by all the children*
Verb	19	14	74
Noun	34	11	32
Adjective	11	0	0
Pronoun	2	1	–*
Prepositions and Adverbs	5	1	–*

Thus the verbs are clearly the most familiar words with the nouns a long way behind. The reason for this may be that the verbs deal with actions which all the children execute, and these words are therefore highly relevant to their lives. The nouns may be more remote, i.e. all children eat and play but some may not be interested in horses. The most remote nouns of all—abstract concepts—are represented only once in this list, by the word 'fun'. This falls in the 'least familiar' group. Edwards and Gibbon have noted that the 'plateau' in reading ability comes at that level in books where the percentage of nouns suddenly increases (Edwards and Gibbon, 1964). They ascribe this to the irrelevance of these nouns to the children's egocentric worlds. The same effect appears to operate here.

Discussion

The scores obtained in this study show that this group of children, having an average level of verbal comprehension, was by no means familiar with all the common words it would soon be asked to read. Such a conclusion could never be reached by studying the results of global measures such as the EPVT, or standardized tests of verbal ability. Average scores obviously conceal individual difficulties and it is these same individual differences that may lead to later difficulties. It has long been a truism, quoted, for example, in the Plowden report (Central Advisory Council for Education, 1967) that a vocabulary of 3,000 words is a necessary pre-requisite for reading. No doubt this group of children, with average scores on the EPVT all had vocabularies of the required size. The point is that these vocabularies lacked critical words.

---

*The number of words in this category is too small for percentages to be valid.

The result also serves to reinforce Widlake's hypothesis that lower SES equals lower vocabulary score. The average EPVT score of our mainly working class group (Mean = 100, $\sigma = 14$) falls neatly between that of his disadvantaged EPA sample (Mean = 93·1, $\sigma = 15{\cdot}3$) and his middle class control group (Mean = 108·0, $\sigma = 7{\cdot}2$). Both differences are significant at the 0·01 level (*t* test). So one would guess that his EPA children would score even lower than the average of 88 per cent of our group obtained on this test. There is a minimal amount of evidence for this speculation in that the middle class children in our sample averaged a score of 91 per cent while the working class children average 87 per cent. However, the sub-samples are too small, and the background data too unreliable to allow any significance to be deduced from this. Were this supposition tested and upheld, there would be a strong case for specific vocabulary building, in addition to general language enrichment, in EPA Pre-School Programmes.

Both the EPVT and our own test are picture recognition tests which assess the child's ability to recognize the representation of an object rather than the object itself. This is an intermediate level of abstraction between the solid object and its representation as a printed word on a page. The child, from a background where the prevailing mode of expression is restricted and concrete, must learn to cope with such pictorial representations before he can grasp the principle of that greater abstraction which is writing. This may explain the class differences in tests of this nature; it also explains why such measures are more relevant to assessment of pre-reading skills than tests involving responses to 'real' objects.

It must be stressed that all the children in the sample were attending good LEA nursery schools where adequate attention was being paid to language enrichment and picture interpretation. Yet this highly-privileged group responded 100 per cent correctly to only 29 of the 73 criterion words. One could posit that children from similar backgrounds without the benefit of nursery education would recognize even fewer words than this.

When we consider the percentage of children familiar with each word we realize again how individual the child's vocabulary is. This is, of course, implicit in the practices of modern nursery education, but is not so obviously in the practices of those writing children's primers. The vocabulary content of these is often based on American word counts which are themselves derived from children's literature, i.e. books written for children by adults. There is no reason at all why English children should know the words American adults think they ought to! There is only one available count of the expressive vocabulary of English school children and that is Edwards and Gibbon's Graded Vocabulary Survey (Edwards and Gibbons, 1964). The writings of 820 young children were analysed and each word was given a popularity index reflecting its frequency of occurrence in these writings.

The youngest children surveyed in the Graded Vocabulary Survey were five-year-olds so the results may not be strictly comparable as vocabularies expand rapidly at this age. Of the 73 words used in our test, seven did not appear in the GVS, i.e. were not used spontaneously by children of this age group. These seven words had a 'knowledge score' of 73 per cent, which is well below the average of 89 per cent for the remainder of the list and shows, not surprisingly, that the children do not know the words they do not use. Of the words known by over 90 per cent of our children, 65 per cent appear in the first 250 spontaneously used by children aged 5+, while of the remainder only 52 per cent are spontaneously used.

The authors of the GVS point out the lack of agreement between the words children use and the content of elementary reading books. We have shown the lack of agreement between the words children understand and these same books. We would join Edwards and Gibbons in their censure of those who would impose on children their own conception of a child's vocabulary.

Conclusion

This study has shown the disparity between the words recognized by pre-school children and the content of elementary reading books. Faced with this problem we have two courses of action—we can either alter the children to fit the books or we can alter the books to fit the children.

Altering the children to fit the books means that nursery school and reception class teachers should realize that although their children appear to be of an average level of verbal ability and seem to possess an adequate vocabulary they may still not know all the words they will find in their new reading books. A check on the child's knowledge of the words in the scheme under consideration would be useful both in assessing his level of reading readiness and in deciding on the nature and content of pre-reading work.

However, this disparity perhaps shows the need for a reappraisal of the content of elementary reading books. The wide differences in the content of our subject's individual vocabularies would seem to indicate the usefulness and relevance of reading programmes based on the child's own vocabulary. It is heartening to note the emergence of such schemes.

References

BRIMER, M. A. and DUNN, L. M. (1965). *The English Picture Vocabulary Tests*. Bristol: Educational Evaluation Enterprises.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (1967). *Children and their Primary Schools* (The Plowden Report). London: HMSO.

EDWARDS, R. P. A. and GIBBON, V. (1964). *Words Your Children Use*. London: Burke.

MENYUK, P. (1963). 'Syntactical structures in the language of children', *Child Develop.*, **34**, 407–22.

MENYUK, P. (1964). 'Syntactical structures used by children from pre-school through first grade', *Child Develop.*, 35, 5, 33–46.

THE REGISTRAR GENERAL (1968). *Statistical Review of England and Wales for* 1966 (*Part II*). London: HMSO.

WIDLAKE, P. (1971). 'Some observations on the language of pre-school children in an educational priority area', *Educ. Rev.*, **24**, 1, 59–69.

Reading schemes used

The Beacon Readers. Ginn & Co. Ltd.
The Downing Readers. i.t.a. Publishing Co.
The Gay Way Series. Macmillan & Co. Ltd.
Happy Venture Readers. Oliver & Boyd Ltd.
Janet and John Reading Course. James Nisbet & Co. Ltd.
The Ladybird Key Words Reading Scheme. Wills & Hepworth Ltd.

Some Determinants of Teacher Morale in Australia

G. M. Coverdale, Macquarie University, Australia

Summary

How satisfied is the Australian teacher with his job and how committed is he to teaching? This paper[1] reports the results of an investigation into the morale within the teaching profession. Using the questionnaire method an attempt was made to uncover factors to account for the high level of resignations and the general dissatisfaction with present conditions of teachers in New South Wales. Conditions of service rather than salary or other material benefits were by far the main concern of respondents who had a wide range of teaching experience and held various positions in both primary and secondary schools. The inspectorial and transfer system, promotion structure and status, class sizes and the demands on the teacher and scope of the curriculum were all ranked as highly important issues affecting morale by a majority of the teachers. Other factors related to morale including the social background of teachers, motives for taking up teaching and reflections on the usefulness of teacher training were also investigated. The article ends with a selection of teachers' comments.

It would hardly be an exaggeration to say that all is far from well in the world of the Australian teacher at the present time, and that morale has plummeted to alarmingly low depths.

In November 1971, the School of Education, Macquarie University, was invited to make a submission to the Senate Standing Committee on Education, Science and the Arts. The opening paragraphs of the Davidson Report (1972) read:

> There is considerable malaise in Australia about education, focussed in the main on teachers and the preparation of teachers, and on what happens within the schools.
>
> Industrial unrest among teachers in more than one State, frequent protests in the press and in the form of petitions and demonstrations, allied with condemnatory statements from senior academics and members of various groups in the community, are symptoms of the general disquiet. Indeed, the very existence of the Senate Committee on the role of the Commonwealth in Teacher Education is itself strong evidence that all is far from well in education in Australia.
>
> It is difficult, however, to point to hard data to support the many claims about the weaknesses of the system as a whole, although there is ample evidence about the deficiencies in various sectors. Little effort has been made to place this evidence in context or to answer the fundamental questions about the provision of education.

In this article I am particularly concerned with job satisfaction, commitment, and general morale within the Australian teaching profession, and in this context, the Macquarie University submission had this to say:

> There are compelling considerations relating to the quality of education, which depend to a great extent on the quality of teachers both in terms of knowledge and skills, and of their dedication and motivation.

Wastage of teachers

There are four critical factors relating to the 'dedication and motivation' issue which should be emphasized, and can be supported by figures from the Bell Report (1971).

1. The drop-out rate of students (particularly university students), undertaking courses of teacher education is singularly high. Twenty per

---

[1] A fuller account of this study in a monograph, 'Teacher Morale: a pilot study', is available from the author.

cent of graduates and eight per cent of non-graduates withdrew from teacher training colleges in New South Wales in 1969.

2. The resignation of newly-qualified teachers serving out their bond reflects serious wastage, which is especially high during the second year of teaching (35 per cent in New South Wales during the period 1968-70). A period of bonding, which varies for graduates and non-graduates from three to five years depending on length of training, is imposed as a condition for the awarding of teachers' scholarships in Australia.

3. Overall teacher resignations (excluding retirements and deaths) (3,822—10 per cent by mid-1971 in New South Wales) tend to exceed recruitments from training establishments (2,072), and, until the recent slight recession, resignation rates have been steadily increasing in State schools throughout most of Australia.

4. In this sort of climate, it is quite possible that for every teacher who resigns, there are many of his colleagues who remain disconsolately in their classrooms in a mood of depression, discontent, and possibly even despair.

The question is: how would one account for this situation? Many would argue that in a relatively affluent society with near full employment, it takes quite an exceptional person to turn a blind eye to the opportunities which exist outside teaching. Others might point to the fact that there are numerous students who, for purely financial reasons, would be unable to attend university if it were not for their teachers' scholarship. They have no intention of taking up teaching as a career and reject it as soon as they can. The third major factor is this critical question of morale.

The concept of morale

The term 'morale' has always proved somewhat elusive to define. It concerns the mental or emotional attitudes of teachers towards the components of their job. It takes into account the atmosphere or 'climate' in which they work, and their individual orientation towards their task. It is essentially a reflection of how one feels about things, and is therefore a matter of subjective perception rather than objective fact.

As a group phenomenon, morale is expressed by:

(a) tenacious persistence and energy in enduring and attempting to overcome difficulty and frustration;

(b) enthusiasm and zealous striving in pursuit of the school's objectives;

(c) group cohesion and co-operative functioning of the teachers who comprise the staff of a school.

Features of low morale are:

(a) a tendency to elevate personal interests above the purpose of the enterprise;

(b) failure to derive personal satisfaction from group achievement;

(c) behaviour that is obstructive and non-contributory to the common purpose.

Clearly, morale implies some human quality which prompts a person to produce at maximum output, and without which he cannot perform at his best. It is associated with a forward-looking, healthy and confident state of mind and includes such attributes as persistence, enthusiasm, zeal and pride. It can usually be increased by favourably modifying any condition that will increase job-satisfaction (Carr, 1970; Gilbam, 1963; McLaren, 1968; and Sergiovanni, 1967). There is a clear relationship between teacher morale and pupil achievement.

The study

In 1971 I conducted a pilot study of a random sample of practising primary and secondary State school teachers throughout New South Wales. This was purely a piece of exploratory research on teacher morale designed to highlight some of the key issues worthy of further investigation in depth as an on-going project, but even as exploratory research, it provided a great deal of information.

The 750 questionnaires distributed, each consisting of 61 questions, some open-ended, were accompanied by a letter which explained the study's purpose:

> It may be argued that much of the contemporary research in the field of education may remain somewhat academic unless its findings can be actively translated into practice through the interest and enthusiasm, the devotion and expertise, of the teaching profession, and that, at this point in time, some investigation is urgently needed into the status, attitudes, grievances and general morale of that profession.

The format of the questionnaire which was arrived at after much discussion with colleagues and teachers included six main sections:

(i) the role and the self-image of the teacher;
(ii) a list of problems raised by teachers to be graded on a four-point scale, ranging from the category of 'highly important' to that 'of little or no importance';
(iii) social background of the teacher and his family;
(iv) retrospective assessment of teacher training;
(v) recommendations by the teachers themselves for improving the profession;
(vi) some personal details of the respondent.

In every case, provision was made for respondents to preserve their anonymity where they so desired.

How representative were the respondents? Every effort was made to ensure a random sample, although it has to be conceded that a certain emphasis was placed on experience. One hundred and sixty-five replies (22 per cent) were received in all. For every respondent with less than ten years of classroom experience, there were four who had been teaching for more than ten years, and half of these, in fact, had been teaching for over twenty years. On the whole this bias favours maturity of judgement, though some might argue that it also gives a voice to those who have grown stale, cynical and disillusioned. In any case, it may be difficult to obtain a truly random sample in an enquiry of this nature. The disenchanted are quite likely to commit their thoughts to paper, whereas the content or the complacent are less likely to respond. For all these reasons the findings of this investigation can only be tentative, although an important supplement to the paucity of existing research on morale.

The proportion of responses from primary school teachers was approximately 25 per cent, whereas the proportion contacted in the sample was approximately 45 per cent (Table 1). This would seem to imply a greater acceptance of existing conditions by primary school teachers, a possibility which is borne out by the resignation figures for those under bond, where the primary schools showed up better than the secondary schools. This would not seem surprising, bearing in mind the relative informality and the more personal atmosphere of the primary school, together with the child-centred approach, the relative absence of curriculum constraints, and the comparative freedom from serious disciplinary problems.

Table 1: Details of respondents according to the level of teaching and the sex of the teacher

	Male	*Female*	*Total*
Primary School	35	7	42
Secondary School	81	42	123
Total	116	49	165

Although women were only 27 per cent of the original sample, 30 per cent of the respondents were women. However, these figures do not accurately reflect the proportion of women teachers in the profession. The positions held by all the respondents and their teaching experience can be seen from Tables 2 and 3.

Teachers' main grievances

The first question involved the ranking, on a four-point scale (highly important, important, of some importance, of little or no importance), of 38 problems repeatedly raised by teachers as affecting their morale. It is interesting to note that salary (and teachers are by no means overpaid by Australian standards), came twentieth in the ranking. The teachers' main concern was with conditions of service and this transcended all other considerations.

(1) *Existing system of inspection*

Three factors were regarded by 66 per cent of the respondents as highly important issues. Most

Table 2: Positions held by respondents

Principal: secondary	4
Principal: primary	16
Deputy Principal: secondary	8
Deputy Principal: primary	7
Subject Master/Mistress: secondary	41
Special Master/Mistress: secondary	14
Infants Mistress	3
Assistant Teachers	72
Total	165

Table 3: Teaching experience

20 years or more	69
15–19 years	28
10–14 years	37
5–9 years	18
less than 5 years	13
Total	165

respondents expressed a deep dissatisfaction with the present inspectorial system, feeling that nothing is calculated to destroy the professional self-image more than a system of close supervision administered so often in such a formal, unbending and frequently humiliating way. It is the way in which inspection has been conducted for so long which caused concern and dissatisfaction rather than the necessity of some form of inspection. Modifications to the inspectorial system are now under way.

Related to this issue was the inhibiting effect of the promotion system which ranked high as a source of discontent (for 54 per cent of respondents). Many apparently enterprising and enthusiastic young teachers resent, and are disheartened by, a promotion system which slavishly favours seniority measured by length of service. It is also debatable whether such a system is in the best interests of the schools. Another factor is whether or not promotion should be synonymous with appointment to positions which are largely administrative. From the responses, it appears that a number of teachers feel that promotion should not necessarily exclude them from the interaction of the classroom. It is also argued that a competent teacher does not always make the best administrator.

Similarly, the arbitrary system of transfers is strongly resented (36 per cent). In Australia, a teacher seldom operates in a deprived area because of the challenge it presents, nor in a country area because he prefers the slower pace of life. He goes to such schools because he is sent there, and this can be particularly upsetting for the beginner. He may, in fact, find himself being transferred as many as five times in ten years, which is unsettling for teachers and pupils alike.

(2) *Large class size and inadequate amenities and equipment*

Class sizes, usually around 35-40 pupils, and inadequate amenities and equipment caused great concern amongst respectively 66 and 54 per cent of respondents. These are factors which reflect a relatively meagre expenditure on education though perhaps this is partly accounted for by the large amount of Government revenue diverted to public works and development projects in a country as new as Australia. It is highly unlikely, especially with the school population steadily increasing, that class sizes will significantly decrease in the foreseeable future. This means that teachers will have to have greater recourse to technological aids in the conduct of their classes if an optimum rate of progress is to be maintained. This makes apparent the need for improved amenities and equipment. Also for the more generous provision of refresher courses so that every teacher may become more fully conversant with the educational media which should be at his disposal in this day and age.

(3) *Intensification of demands on teacher and rigidity of curriculum*

Another area of discontent centres around the rigidity of the curriculum, and the way in which the teacher has been dictated a functionary role aimed at getting as many children through the system as is practicable at the highest levels of achievement possible. The result is that so much of the challenge and idealism of teaching, and the creative opportunities open to teachers, are lost in the process. Under these circumstances, teaching is no longer a distinctive occupation. Teachers are relegated to a utilitarian role with little or no say in policy-making and expected to concern themselves with classroom matters only. It is not surprising if teaching fails to attract the sharpest and most searching minds among the young, especially with the added discouragement of lack of prestige, occupational status and professional recognition. Fifty-four per cent of respondents thought that the responsibilities of the teacher were not really recognized as a profession by the public.

It is appreciated that lack of prestige and professional recognition is also very much an issue in most Western countries, unlike the USSR where teaching is a prestigious occupation and the teacher is a very potent instrument of 'the Party'. In a dynamic, striving and materialistic society like Australia, the teacher is all too often thought of as a mediocre practitioner of average intelligence but without the initiative or drive to enter a more rewarding or demanding career. The curriculum does not help his self-expression as it is usually so lacking in relevance and motivation, besides being virtually impervious to any depth of innovation.

Linked with status was an expression of disquiet by 54 per cent at the absence of an independent Education Commission which might serve to arbitrate in matters of professional dispute between teachers and their very remote employers. It might also give a voice to those who, classed as public servants, are virtually denied the normal rights of freedom of speech.

Two other prominent issues were highly ranked by 41 per cent of the respondents. One concerns what the teacher sees as the indifference of parents in a country where community involvement and lay participation in education are at a notably low level. The other is the problem of immigrant children with language difficulties.

From the responses to this question, these then appeared as the major issues affecting teacher morale at the present time. Many of them it must be conceded, are the sole province of the employing authorities. It is only their wisdom and foresight which can alleviate the concern which the teachers feel for their working conditions. There are some signs of hope for the future: a belated injection of finance into the education system, proposed modifications to the inspectorial system, less slavish adherence to external examinations, and a somewhat latent recognition that a higher degree of effective decentralization would hopefully lead to better human relations.

One might argue that it is also up to the universities to play their part and help to elevate the profession to its rightful status in every way they can, with especial reference to inculcating a greater degree of contentment and stability into the practitioners of the future.

Other factors related to morale

The fact that there is little 'continuity of occupation' in the teaching profession was firmly established by a question enquiring into parental and grandparental occupations. Out of a possible total of nearly a thousand, only 30 immediate antecedants had been teachers (see Table 4). Just over 40 per cent of the teachers were from the homes of manual workers. This gives credence to the concept of teaching as a 'bridge profession' providing a means of 'upward social mobility' rather than establishing a family tradition and an enduring commitment to the profession.

Linked with this question was one which enquired into the motives for entering the teaching profession, many respondents giving more than one reason. Thirty-five per cent claimed to do so through a sense of vocation. Twenty-two per cent apparently gravitated towards teaching because of the lack of alternative opportunities, which is a somewhat chastening statistic, especially for those concerned with the selection of the teachers of tomorrow! Another 22 per cent chose teaching because of the opportunities they hoped would be offered for creativity and initiative and for a challenging and satisfying career. Other factors were security (20 per cent) and salary and working conditions (20 per cent), the latter being particularly important to married women. Only two per cent mentioned prestige!

Table 4: Details of fathers' and grandfathers' occupations expressed as percentages

	Fathers	*Grandfathers*
Professional/semi-professional	9·2	7·6
Administrative/proprietary/ managerial	21·9	15·2
Teachers	6·7	3·1
Clerical/Sales	7·3	6·0
Rural enterprise	13·5	27·6
Skilled manual	25·1	23·4
Unskilled manual	16·3	17·1
Total	100·0	100·0

Asked whether they would encourage their children to become teachers, 22 per cent gave an unqualified 'yes'; 26 per cent an unqualified 'no'; and 16 per cent said 'yes' for daughters and 'no' for sons. The remainder gave answers which were either very guarded or else heavily qualified. Their main anxieties included the public's attitude of complacency towards education, and the teacher's lack of professional and social status and an absence of job satisfaction.

Asked to evaluate their teacher training in retrospect (60 per cent were graduates), two-thirds of the respondents regarded it as 'inadequate'. Many used stronger terms such as 'useless' and an 'utter waste of time'. The dissatisfaction expressed with teacher training came equally from teachers who had been fully trained in teachers' colleges and those who had completed a postgraduate Diploma of Education. Only one teacher in four considered that his training had been 'adequate', 'generally good' or 'helpful'. A major criticism was that too much emphasis had been placed on content and presentation and not enough on a sociological understanding of the child.

Many respondents called for a more practical and down-to-earth approach to matters such as group behaviour, management and discipline, methods of motivating pupils, and the teaching of low ability and culturally deprived children. It was felt that such topics were often overlooked or else treated in a theoretical and academic manner. In this context, it was suggested that lecturers should return regularly to schools to up-date their teaching experience. Forty per cent of the respondents stressed the need for longer and more intensive periods of practice teaching.

Strong support was given to the idea of regular in-service refresher courses. Almost 75 per cent of the respondents considered them necessary. One-third of these thought that they should be compulsory. There was a strong feeling that such courses, by bringing teachers together to share ideas and discuss common problems, would be a means of boosting knowledge and increasing morale. Other suggestions for improving and stabilizing the teaching profession correlated closely with the problems enumerated earlier as causing most concern and thus influencing morale.

Teacher comment

Finally, a few miscellaneous quotations extracted from teachers' responses to the questionnaire.

(i) The dedicated pupil-centred teacher

My aim is to help create happy, healthy children who can think for themselves. I try to teach them not what to think but how to think. Within the limits of a stifling curriculum I try to encourage them to question all that is put before them, to accept no one's word at face value, to listen to advice, yes, but to work things out for themselves. I want to continue doing this as a classroom teacher; I want no promotion.

(ii) The realistic woman teacher

Disenchantment? At our school nearly all the staff moan, but nobody wants to go. The men are much more disenchanted than the women. Men suffer from the belief that there is somewhere in life to get, and that they will find happiness once they get there. Women are much more realistic. No doubt this is the same in other jobs.

(iii) The disillusioned missionary

First let me say that I enjoy teaching very much and intend to make it a career. I do feel, however, that I am fighting considerable odds to maintain my enthusiasm. I think these odds will overwhelm me and cause me to resign out of frustration. Not only am I fighting arbitrariness and conservatism from my employers but utter indifference from parents.

(iv) The inveterate pessimist

My present role is a depressing round of making fresh allocations and timetables as staff continuously come and go (mostly go); of attempting in very limited time to help the younger members professionally (planning lessons, using equipment, suggesting approaches); of trying to keep going an antiquated system with decaying classrooms, gross overcrowding, insufficient staff, rampant authoritarianism, and a dearth of ideas.

(v) And to conclude on an optimistic note

I do see some hope for the future. A free approach is in the air, restrictive examinations are slowly going, more experimentation is occurring. I like to think, and this is what keeps me and the majority of this staff going, that within ten years we will see all exams abolished, self-discipline rather than authoritarianism in schools, a public aware of what education really means and not being satisfied with the narrow vocational system they have now, and a more enlightened administration. Perhaps I am a super-optimist, but the people with ideas and energy are in the schools—they simply need a freer hand.

References

BELL REPORT (1971). *Report of the Committee of Enquiry into Teacher Education in New South Wales.* Sydney: Department of Education, New South Wales.

CARR, D. (1970). 'What is wrong with the teaching service?', *Quadrant*, **14**, 57–62.

DAVIDSON REPORT (1972). *Report on the Commonwealth Role in Teacher Education.* Canberra: Senate Standing Committee on Education, Science, and the Arts.

GILBAM, R. J. (1963). 'Teacher attitudes towards some administrative practices in a sample of N.S.W. State High Schools', *Journal of Educational Administration* (University of New England), **1**.

MCLAREN, J. (1968). *Our Troubled Schools.* Melbourne: Cheshire.

SERGIOVANNI, T. (1967). 'Factors which affect satisfaction and dissatisfaction of teachers', *Journal of Educational Administration*, **5**, 66–82.

Authoritarian Personality Traits among Students

J. Dunham, University of Bath

Summary

The results are reported of an investigation into the social attitudes of the total population (800) of one English university using Adorno's F scale to measure authoritarian personality traits. Analysed by departments it appears that more scientists are ranked in the 'authoritarian' group (over 50 per cent were engineers and 25 per cent mathematicians and pharmacists) whereas more arts students fall into the non-authoritarian band (over 30 per cent sociologists, and 20 per cent architecture students). The author considers possible objections to the test and compares his own results with a review of cross-cultural and international studies in the field. Other tests of authoritarianism are discussed and evaluated.

Social attitudes of students and the authoritarian F scale

During the past forty years the attitudes of students towards such issues as war, communism, trade unions and religion have frequently been investigated. There is evidence that students in certain fields of study have consistently tended to have conservative social attitudes and equally consistently students in other fields have tended to hold liberal attitudes. Usually, but not always, in these investigations, social science students come out as the most liberal of the groups. There is much more certainty that students in engineering and agriculture will be found in the least liberal groups.

One of the most frequently researched topics in the field of social attitudes has been 'authoritarianism'. This, until fairly recently, has usually been assessed by means of the F (for Fascism) scale, which was prepared by Adorno *et al.* (1950). This questionnaire was based on their concept of the characteristics an authoritarian person was likely to have. According to these investigators, an individual with strongly developed authoritarian traits 'rigidly adheres to conventional middle-class values and has an exaggerated concern with such values, is submissive toward the moral authorities of his in-group, condemns and rejects people who violate conventional values, is preoccupied with power and status considerations, tends to identify with powerful figures and is generally hostile towards members of out-groups'. This scale consists of statements which represent widely held opinions on various social questions, e.g.

'Familiarity breeds contempt'

'Obedience and respect for authority are the most important virtues children should learn'

'Young people sometimes get rebellious ideas but as they grow up they ought to get over them and settle down'

'People can be divided into two distinct classes; the weak and the strong'.

When the scale is administered, respondents are asked to indicate their opinions on each of these statements on a seven point scale ranging from strongly agree, through neutral to strongly disagree. The higher the person's score the more authoritarian he is said to be. The question for investigation in this section is whether students in different faculties in universities in America and England are consistently different in their scores on the F scale and other tests of authoritarianism.

Stern (1961) has reported studies using the F scale conducted at the Centre for the Study of Higher Education at the University of California. Freshmen who had decided to read engineering made higher scores, i.e. were more authoritarian, than freshmen selecting mathematics, physics, education and medicine. Dental students were 'authoritarian and inflexible' with marked conformity needs. On the comparisons made by means of scores on personality inventories, students in engineering and dentistry appeared to be most alike. Social science students on the other hand consistently made low, i.e. non-authoritarian scores.

The investigation

In investigations which I conducted at an English university the F scale was first administered, in a pilot study, to 52 first-year students; 26 of them were in an architecture class, 11 were in an engineering class and 15 were reading sociology. When their F scale scores were ranked, the top quarter were placed in the authoritarian group. The mean of the authoritarian group's score was 124 and the range was from 101 to 163. The mean of the non-authoritarian group was 65 and the range was from 45 to 78. The subject composition of the groups was in line with research at the University of California: 63 per cent of the engineers were in the authoritarian group and none in the non-authoritarian, 14 per cent of the sociologists were authoritarians and 35 per cent non-authoritarian, 46 per cent of the architects were authoritarians and 20 per cent non-authoritarians.

This pilot study, which had indicated differences in authoritarianism and opinions about relationships with authority figures between first year engineering, architecture and sociology university students was followed up a year later by an investigation which involved the total university population of nearly eight hundred students. When the F scale scores of these students had been scored and ranked, the top sixth were placed in the authoritarian group and the bottom sixth in the non-authoritarian group. There were approximately 130 students in each group.

Results and discussion

In the authoritarian sample over half were engineers and a quarter were mathematicians and pharmacists. In the non-authoritarian sample almost one third were sociologists and one-fifth were architecture students.

The actual numbers and percentages in each department in the (most) authoritarian sample were:

Electrical Engineering	30%	$n=21$	Total $n=$ 71
Mathematics	25%	$n=12$	,, $n=$ 48
Engineering	23%	$n=49$	,, $n=214$
Pharmacy	20%	$n=15$	,, $n=$ 75
Architecture	16%	$n=14$	,, $n=$ 87
Economics & Administration	14%	$n=$ 4	,, $n=$ 28
Statistics	13%	$n=$ 2	,, $n=$ 16
Chemistry	11%	$n=$ 9	,, $n=$ 79
Applied Biology	10%	$n=$ 7	,, $n=$ 70
Applied Bio-Chemistry	8%	$n=$ 1	,, $n=$ 12
Applied Physics	8%	$n=$ 2	,, $n=$ 26
Sociology	1%	$n=$ 1	,, $n=$ 62
Building Technology	0%	$n=$ 0	,, $n=$ 10

The lone authoritarian sociologist called the writer 'Sir' for four years. He thought that 'Formal pupil-teacher relationships are necessary for some students even in their final year at university' and agreed that students ought to be out of bedrooms occupied by the opposite sex or else 'Brothel-like conditions would prevail'. He did his sandwich placement away from the university in the Police Force and eventually went to work in an Approved School.

The numbers and percentages of students in each department in the non-authoritarian sample were:

Sociology	66%	$n=41$
Applied Physics	31%	$n=$ 8
Architecture	30%	$n=26$
Building Technology	30%	$n=$ 3
Economics & Administration	25%	$n=$ 7
Statistics	19%	$n=$ 3
Chemistry	11%	$n=$ 9
Applied Biology	10%	$n=$ 7
Mathematics	10%	$n=$ 5
Applied Bio-Chemistry	8%	$n=$ 1
Engineering	7%	$n=15$
Electrical Engineering	4%	$n=$ 3
Pharmacy	3%	$n=$ 2

These results are in substantial agreement with the pilot study and with the American investigations discussed earlier. As before, engineers appeared to be tending to score in the authoritarian direction and sociologists in the non-authoritarian. But this group of mathematicians appeared to have more inflexible, intolerant and conventional attitudes than those in the University of California Study. As in the pilot study the architects appeared to have the widest range of F scale scores indicating marked differences in their attitudes. Economics and Administration students also showed wide variations.

Other studies using the F scale

In a similar study at another English University, Lee (1968) tested the hypothesis that 'There is a relationship between degree of authoritarianism and the students' university faculty'. By using first- and third-year students she was also able to investigate the possibility that these personality characteristics may change during the university course. Her testing with the F scale led to the following results:

1. The F scale scores of the Science students were significantly higher than those of the Arts students.
2. The Engineering students scored significantly higher than the Arts students.
3. The Engineering students made significantly higher scores than the Social Science students.
4. The scores of the Medical and the Arts students were not significantly different.
5. The scores of the Science and Social Science students were also not significantly different.

When she compared the first- and third-year students in engineering and medicine she found that in both faculties the third-year students had made significantly lower scores.

Because these results and those from my own investigations provide evidence of student attitude differences in university faculties it is important to note the criticisms, which have been made of the F scale. Its statements are said to be ambiguous and therefore subject to different interpretations. It is said to indicate only authoritarianism of the Right, i.e. Fascism, and to be misleading about extreme Left-Wingers, who by scoring low appear to be non-authoritarian. Thirdly, it has been strongly criticized because high scores do not always indicate a high level of intolerance, rigidity and prejudice. They might also be the result of 'response set' in the individual completing the questionnaire. Some people might 'agree' with the first few statements and continue this pattern of response, because in the scale there are no negative statements. So high scorers may be people who answer questionnaires in this way (Hyman and Sheatsley, 1954).

Though these criticisms warrant caution in the interpretation of the test, the F scale continues to be used for the assessment of authoritarianism in contemporary social psychology. In one recent study, Meade and Whittaker (1967) used the F scale in English, Hindi and Portuguese to test American, Arabian, Chinese, African, Indian and Brazilian students. The results indicated that the Indian and African students were the most authoritarian students. The least authoritarian were the Americans and Brazilians.

In their discussion of these results, the two investigators consider the origins of authoritarianism. They say, 'Differences in religion, family structure, child-rearing practices, education and type of government, working singly or in interaction with others, may provide an explanation. The cultural groups, which are represented here quite obviously differ with respect to all of these factors and many more as well. Further research is needed before specific factors contributing to authoritarianism or lack of it in each culture studied are isolated'.

Other tests of authoritarianism

The criticisms, which were made of the F scale, resulted in a reduction in the number of investigators using it and the development of other questionnaires. One of these was produced by Stern and his colleagues at the University of Chicago and named by him, an 'Inventory of Beliefs' (Stern, 1961).

The Inventory includes the following statements:

1. 'Reviewers and critics of art, music and literature decide what they like and then force their tastes on the public'
2. 'Science is infringing upon religion when it attempts to delve into the origin of life itself'
3. 'A lot of teachers these days have radical ideas which need to be carefully watched'
4. 'We would be better off if people would talk less and work more'.

The individual who agrees with the statements, is thought of as 'immature, rigid in his outlook, authoritarian in his relationships with others and compulsive in many of his actions.' The person who disagrees with the statements is thought of as adaptive, flexible, mature and democratic in his relationships with other people. In the scale disagreement with the statements is marked high numerically.

Studies using the Inventory which have been conducted in American universities and colleges reveal significant differences in occupational choices between high and low scorers. 'The high scorers prefer occupations involving interpersonal, expressive or abstract activity, such as psychology, music or theoretical physics. The low scorers, i.e. those who generally agree with the statements prefer more impersonal, concrete or status-orientated occupations such as engineering, medicine or law' (Stern, 1961).

A much more convenient technique for the assessment of authoritarianism has been developed recently by Wilson and Patterson. They argued that there is 'a variable of personality, which has been shown to have considerable predictive value in both social and clinical psychology and which has been variously labelled authoritarianism, dogmatism, fascism, and anti-scientific attitude'. These investigators are critical of this label, which is 'value-toned', and they propose to use 'conservatism'. They are equally critical of previous attempts, including the F scale, to measure this variable. Instead they offer 'A New Measure of Conservatism'. In their questionnaire Wilson and Patterson present 'a list of brief labels or catch-phrases representing various familiar and controversial issues'. The respondent is asked, 'Which of the following do you favour or believe in?'

1. Death penalty
2. Evolution theory
3. School uniforms
4. Striptease shows
5. Sabbath Observance, etc.

It is too early yet to report more than a few investigations recording the performance of various groups on the 'C-Scale'. But those that have been reported by Wilson and Patterson (1970) indicate support for the hypothesis that students in different faculties can be differentiated in this personality syndrome.

Some of the norms given are:
University Students (UK) — Mean = 25.33
College of Education
students (UK) — Mean = 36.07
College of Education,
Women students (UK) — Mean = 43.10

Education students in college are on this test more conservative than non-education university students. But the two scores for the college students are even more important for the argument as the women in the sample were students in a College of Home Economics. Their higher average score indicates, that generally they are more in favour, even than students in other colleges, of such 'conservative' characteristics as 'religious dogmatism, right-wing political orientation and insistence on strict rules and punishments'.

Discussion and evaluation

The research investigations examined so far in this paper have indicated that students in various fields of study have different social attitudes. These studies are inconclusive for several reasons. The number of subjects was so small as to make them possibly unrepresentative of the larger bodies of students they are meant to sample. They may not be typical because there is little evidence that, for instance, engineering students in one university are similar to those in another. Similarly, Social Science students are following a wide range of courses with considerable variation in subject matter. Within each course there is a wide range of scores. Between the courses there is considerable overlap. But there does appear to be some evidence that there are differences in social attitudes which merit the attention of teachers responsible for the courses.

This conclusion is strongly reinforced by the evidence that the effectiveness of different university environments varies with the social attitudes of their students. This indicates that students who have a strong need for direction and organization will perform best in structured and formal learning situations while others, who place a high value on personal independence, will enjoy and perform well with less structured, informal and non-directive teaching techniques. One research worker, Stern (1962) has reported that those students with the first group of characteristics in the University of Chicago did unusually poorly in the social sciences and the humanities and strongly disliked these subjects.

In his investigations of the interaction between student characteristics and the 'press' of college environments Stern assessed the way students perceive themselves and the characteristics of their colleges by means of the Student Activities Index and the College Characteristics Index. The use of these may be shown by discussion of the way students see 'orderliness' in themselves and in their college. Stern defines this as 'a prevailing trend towards the compulsive organization of the immediate physical environment, manifested in a pre-occupation with neatness, orderliness, arrangement and meticulous attention to detail'. The extent of a person's need for 'orderliness' is assessed by the number of preferences he makes in his Student Activities Index for items such as:

'Washing and polishing things like a car, silverware or furniture'

'Keeping an accurate record of the money I spend'

'Arranging my clothes neatly before going to bed'.

The extent of the press in a college environment for 'orderliness' is assessed from the number of respondents from the same institution who agree with such statements in the College Characteristics Index as:

'Attendance is usually taken in each class'

'Student papers and reports must be neat'.

The approach to learning of those students with strongly developed needs for 'orderliness' has been described by Stern as 'somewhat obsessive and compulsive, stressing detailed organization and structure'. Their attitude to university education favoured vocationally orientated courses. When the academic environment they experienced at Chicago University was very different from their expectations they became so dissatisfied that they 'dropped-out'. They gave two major reasons for dropping out. First they complained about the lack of professional courses. Secondly they criticized their lecturers because they did not make attendance at lectures compulsory, and allowed smoking in the classrooms.

In a study at Harvard University, Jacob (1957) reported that, 'permissive type teaching caused some students to experience great emotional strain. They were confused, even angry; whereas under straight lecture and recitation they moved along contentedly'. It appears then that authoritarian students are most critical of non-directive teaching techniques, and that in certain institutional settings they perform poorly because of the lack of congruence these environments have for them.

In his research at the Massachusetts Institute of Technology Parlett made a case-study of one student who was very successful. This physicist knew very clearly the academic environments he preferred. His classes should be 'fairly regimented —not so much making people show up to the class, but having the work done, passed in and graded. In a class that's regimented like that, I learn more'. It was important for him to have examinations because 'If you don't have them, people aren't going to study. I know that I wouldn't. The average person is lazy.' He was critical of one lecturer at MIT because he made errors and forgot things and was not afraid to tell his classes. The student commented, 'This is what's wrong with this class. I like to see teachers in complete command of the situation' (Parlett, 1970).

He appears to be a 'syllabus-bound' student. This means, according to Parlett that he can only learn effectively within the constraints of a syllabus. If he has to choose how and what to learn he becomes less successful. Parlett has developed a Study Habits Inventory (Parlett, 1970). The syllabus-bound student responds to the statements in this inventory in characteristic ways. He is, for instance, likely to answer 'False' to these statements:

'I have several well-developed intellectual interests outside my school work'

2. 'When studying the subject I like most, I find I work much harder'

He is also likely to answer 'True' to these statements:

3. 'Without the stimulus of exams, I doubt whether I would do much effective studying'

4. 'When actually in the exam room my anxiety gets less and I become more efficient responding to the exam'.

The 'syllabus-free' student would respond to these statements in significantly different ways. He also feels most satisfied when pursuing his own work and course structure. He is frequently attracted by and urgently wanting to read books outside the syllabus. He is likely to answer 'False' to these statements in the Study Habits Inventory:

1. 'I prefer not to be left to work in my own way as I get nothing done'

2. 'I usually only study what I am required to study'

He is likely to answer 'True' to these statements:

3. 'I spend too long on certain topics because I get very involved in them'

4. 'I like to play around with certain ideas of my own, even if they do not come to much'.

The academic environment is not however restricted to the syllabus and the book list. It includes something as intangible as 'ethos'. The importance of this is suggested in a study by two investigators who teach in a polytechnic. They found a high degree of conservatism on social and political issues among the engineering and sociology students, 'but the engineers swung further right'. They also reported that these students appeared to be segregated from each other. Differences in inter-group perception were recorded. The engineers saw themselves as 'experienced, practical and hard-headed' and the sociology students as 'immature, idealistic, long-haired and lazy'. The engineers were perceived by the sociologists as 'limited, narrow-minded and thick' (Payne and Bird, 1969).

Is there anything more to these inter-group differences in perception than stereotyping? It may be that they represent the expression of the out-group and in-group comparisons, which groups make to emphasize their own identity. The differences between them may be more apparent than real. The differences within the groups may be larger than those apparently existing between them. All these qualifying objections may be valid. On the other hand, if the similarities only mask the significant differences, certain important implications for teaching would have to be considered. There may really be an engineering sub-culture in a polytechnic or university. If so, it would have norms which are the shared assumptions and expectations of its members. An appropriate system of values, i.e. the 'ethos', would be an important aspect of the way members of the sub-culture related to each other. Any person brought into the sub-culture would be subject to these norms and values. Students who wanted to become engineers and the staff who taught them would be influenced by them. This can be seen quite clearly in the experience of one student whose values, appearance and beliefs were very different from his group norms.

A case study

The student was on an engineering course. He was interested in architecture and liked to listen to poetry written and read by other students. The students on his course thought him odd and peculiar. He was isolated within the sub-culture. His loneliness led to depression and he withdrew further from other people. He stopped going to lectures because they made him more anxious and unhappy. His motivation for engineering became very weak because if the students on his course were like engineers he did not want to be one. Unlike the normative values accepted by the sub-culture, he wanted more from life than obtaining a degree, a good job, getting married, buying a house and owning a car. His hair was much longer than anyone else's on the course and he disliked wearing a suit. He talked of transferring to architecture and to sociology when he came to discuss his difficulties with me in my role of University Counsellor. With counselling he broadened his career spectrum to include design engineering. He began to understand the reality of his own attitudes, experience, qualifications and aspirations compared with those of the sociology and architecture students. He found he was more of an engineer than an architect or a sociologist. In his end-of-year examinations he failed one paper but was allowed a re-sit. This he has since passed. His adjustment to his course may, in spite of his success, never be more than marginal, because the sub-cultural 'ethos' in which it functions, may not change quickly enough to allow him to develop a sense of belonging. Similarly, his own attitudes may continue to prevent a more effective interaction between himself, his group and his teachers.

It appears then that it would not be unreasonable to conclude, first, that students' social attitudes determine to a considerable extent their response to teaching, and secondly, that enough is known about their attitudes to provide students with favourable learning environments.

References

ADORNO, T. W. *et al.* (1950). *The Authoritarian Personality*. New York: Harper & Row.

HYMAN, H. H. and SHEATSLEY, P. B. (1954). 'The authoritarian personality: a methodological critique', in CHRISTIE, R. and JAHODA, M. *Studies in the Scope and Method of the Authoritarian Personality*. Chicago, Illinois: Free Press of Glencoe.

JACOB, P. (1957). *Changing Values in College*. New York: Harper & Row.

LEE, F. (1968). 'The relationship between the degree of authoritarianism and the faculty'. Unpublished dissertation, University of Bristol.

MEADE, D. R. and WHITTAKER, J. O. (1967). 'A cross-cultural study of authoritarianism', *J. Soc. Psychol.*, 72, 3-7.

PARLETT, M. R. (1970). 'The syllabus-bound student', in HUDSON, L. *The Ecology of Human Intelligence*. Harmondsworth: Penguin.

PAYNE, G. and BIRD, P. (1969). 'What are their students like?' *New Society*, 23rd October.

STERN, G. G. (1961) 'The measurement of psychological characteristics of students and their learning environments', in MESSICK, S. and ROSS, J. *Personality Measurement*. New York: Wiley.

STERN, G. G. (1962). 'Environments for Learning', in SANDFORD, R. (Ed.) *The American College*. New York: Wiley.

WILSON, G. and PATTERSON, J. (1970). *Manual for The Conservatism Scale*. Slough: NFER.

Commitment to Teaching of Postgraduate and College-trained Students

David Hellawell, City of Birmingham College of Education
and Alan Smithers, School of Research in Education, University of Bradford

Summary

The commitment to teaching of students in a university department of education, graduates in two colleges of education and students on BEd and three-year courses in the two colleges of education is compared. A higher proportion of those in postgraduate training than on BEd and certificate courses indicated teaching as their ideal career. However, when firmness of intention to make a career in education and hoped-for life-styles were considered, it was the concurrently trained BEd students together with the consecutively trained postgraduates who emerged as the more committed. Thus from this point of view course pattern *per se* does not seem important. It is suggested that commitment to teaching in the sense of intentions rather than aspirations is related to the extent to which career prospects are thought to be more or less favourable. Women students were consistently more favourably disposed towards actual school teaching than men. Implications for the proposals of the James Committee are discussed.

Introduction

The teaching profession is saddled with a poor image. A depressingly large number of people seem to see some truth in George Bernard Shaw's dictum—'he who can does, he who can't teaches'. Even sociologists seem to doubt whether teaching can properly be regarded as a profession (Kelsall and Kelsall, 1970; Leggatt, 1970).

Teaching does carry higher prestige for women than men, and there are important differences within education (this is nicely illustrated by Mary Warnock's account (1970) of her experiences on moving from an Oxford fellowship to a school post). But overall the social standing of the teacher is not high.

This has its effects on the morale of practising teachers and on recruitment. In Rudd and Wiseman's (1962) study, a major source of dissatisfaction, particularly for the male teachers, was the status of the profession. Kelsall, Poole and Kuhn (1970) have interpreted some of the results of their national follow-up survey of the graduates of 1960 to argue that school teaching 'is the choice of people who have, or think they have, restricted opportunities'. A recent report by the Assistant Masters Association (1970) commented that 'too lazy to seek other employment', 'fear of the outside commercial world' and 'too firmly entrenched in the rut of a classroom world' could be taken as indicative of 'the views of many students about their contemporaries who decide to enter the profession'.

Evidence has accumulated over the years that some students are attracted to teacher training for reasons unconnected with any real desire to teach. In the days of Board of Education grants, agreeing to become a teacher was not an uncommon way of financing university studies. Valentine (1934), reporting the results of his inquiry among graduate students in the Education Departments of four universities, remarked that 'nearly all the men who entered an Education Department on entering the university give the Board of Education grant as one of their motives for choosing the profession'.

It is not surprising, then, to find in Evans' (1946) study that students at a Leicester teachers' training college had a significantly more favourable attitude to teaching than the students in the university department with whom she compared them. She attributes this difference to those university students 'who are prepared to teach for a time, then make use of their degrees in other ways. The two-year training, on the other hand, is a preparation for teaching and is of little value for any other career. It therefore attracts only those who are interested in teaching'.

However, with the coming of readily available grants for university education and the ever-increasing pressure on university places, this has changed. In a study in two northern colleges of education, Smithers and Carlisle (1970) found that over half the students would ideally have preferred a career other than teaching. Mangan (1970) replicated the study and found that, even in Scotland, where there is less competition for university places, about a third of his college of education sample appeared to be reluctant recruits to teaching.

Thus the enthusiasm for teaching among those in training, is, in part, a function of what is happening elsewhere. In an attempt to assess the present situation, the authors have recently carried out a survey among students following different routes into teaching. This paper attempts a comparison of their relative occupational commitment.

Methods

The survey was carried out in two colleges of education and one university department belonging to the same Area Training Organization and situated close to one another. Four groups of students were considered: (i) graduates in a university department of education, (ii) graduates taking their postgraduate training in colleges of education, (iii) students in the fourth and final year of BEd courses and (iv) students in the third year of their college of education courses. Of this final group only half had registered for the BEd degree and not all of these were expected to progress to the fourth year.

Both the colleges had fairly large postgraduate departments specializing in training for the secondary level, and the third- and fourth-year college samples were mainly drawn from among those training for this level. The graduate samples did not differ significantly in degree attainment. Seventy-two per cent of the college students had at least two 'A' levels, but the grades were generally poor—only 9·4 per cent having grades better than two Cs.

The study was undertaken in the academic year 1970-71. Questionnaires were completed anonymously during classes (attempts were made to follow-up absentees) at a time when all the students had experienced at least one major teaching practice. The study was small-scale and intensive, but a high degree of co-operation was obtained. Actual numbers and response rates were as follows: (i) graduates in university department of education—100 (94·3 per cent), (ii) graduates in colleges of education—85 (88·5 per cent), (iii) BEd students—41 (89·2 per cent) and (iv) third-year college students—54 (90·0 per cent). The proportions of male and female students in these samples are indicated in Table 1.

The students' attitudes to teaching as a career were studied using three approaches. Following Smithers and Carlisle (1970), students were asked to answer the following question: 'if you were quite free to choose and could obtain the qualifications necessary, what field of employment would you *ideally* like to enter'. They were also asked to indicate on a four-point scale the firmness of their intention to take up teaching on qualifying.

Another approach was taken from Musgrove (1969). Students were asked to 'imagine you have reached the age of 70 and are looking back over your life. What are the highlights and achievements you hope to look back on? List these highlights and achievements with the age at which you hope they would occur.'

Table 1: Intention to teach

	Per cent expressing firm intention:		
Type of course	Males	Females	All
UDE	46·5 (n=43)	57·9 (n=57)	53.0 (n=100)
Colleges— postgraduates	64·7 (n=51)	68·8 (n=32)	66·8 (n=83)
Colleges— BEd	75·0 (n=28)	81·0 (n=11)	76·9 (n=39)
Colleges— 3rd year	44·1 (n=34)	66·7 (n=18)	51·9 (n=52)

n=base

Results

Students' ideal career choices categorized as 'school teaching', 'other education'—usually lecturing in higher or further education, more rarely educational administration—and 'outside education' are presented in Table 2.

Excluded from this part of the analysis are five students who gave research as their ideal career, but were insufficiently explicit for this to be categorized with any confidence (e.g. 'research in optical physics'), ten students who gave two ideal careers, one of which was in education and the other outside (e.g. 'medicine and school teaching') and 16 students who did not respond to this question. Thus the data in Table 2 may tend to over-estimate the proportion of choices given to education.

Among the third-and fourth-year college of education students, 51·2 per cent indicated that their ideal career would have been outside education—a figure very close to that obtained by Smithers and Carlisle (1970) in their study. The third-year students were especially likely to aspire to something else. The postgraduate students in the colleges appeared to be the most satisfied with the prospect of a career in education.

Considering men and women together, and combining the categories 'school teaching' and 'other education', the rank order of ideal choices given to education was: (1) postgraduates in the colleges—71·0 per cent, (2) postgraduates in the university department—62·5 per cent, (3) BEd students—59·0 per cent and (4) third-year college students—39·6 per cent, with the differences between the third-year students and the postgraduates being statistically significant. There was a tendency for female students to be less likely to aspire to a career outside education and the predominance of women in the university department contributes to its high ranking. There was also a tendency for women to be more likely to indicate 'school teaching' as opposed to 'other education' than the men, but the difference just falls short of an acceptable level of statistical significance.

Table 2: Ideal career choices

Type of course	*Per cent indicating:*		
	School teaching	Other education	Outside education
UDE			
m (n=37)	40·5	18·9	40·5
f (n=51)	51·0	13·7	35·3
Colleges—postgraduates			
m (n=50)	42·0	28·0	30·0
f (n=29)	44·8	27·6	27·6
Colleges—BEd			
m (n=28)	28·6	25·0	46·5
f (n=11)	54·6	18·2	27·2
Colleges—3rd year			
m (n=29)	10·3	20·6	69·0
f (n=14)	57·1	nil	42·8

A rich variety of occupations outside education was mentioned with medicine and other élite professions, business, and social work being prominent. Some of the ideal occupations—writer, orchestral work, soccer star—could conceivably be combined, for a time at least, with a career in education. But, on the whole, an impression of making do with second best emerges. Very few of the ideal careers indicated were wild or fantastic. Even 'laser beam operator' from a physics graduate appears to have a serious idea behind it.

The data of Table 2 suggest that the postgraduates as a whole were the more likely to aspire to teaching, but this does not necessarily mean that they were the more likely to make it their career. In Table 1 firmness of intention to enter teaching is examined. The uncertainty of the third-year students is confirmed, but on this measure the BEd students emerged as the most firmly committed to a career in education. Postgraduates in the colleges also tended to be sure that they wanted to take up teaching, but the UDE group were more doubtful. Differences between the BEd students and both the third-year students and the UDE group were statistically significant.

The findings presented in Table 1 are supported by the analysis of replies to the highlights questionnaire. In Table 3 responses have been categorized according to the highest point in the career mentioned. In some cases school teaching has been inferred even if not explicitly stated (e.g. 'to achieve a post of some responsibility'), but 20·4 per cent of the students either did not mention a career at all or said so little that it was not possible to be sure what they had in mind (e.g. 'work abroad, give up smoking, occupational contentment'). Although the proportions of men and women in this category were not very different, there were differences in emphasis. Among their hoped-for life's highlights the female students tended to stress marriage and family (e.g. 'I am totally unambitious and if I can look back on life with the absolute minimum of boredom, with my children and grandchildren happily established, that will be quite enough'), whereas the men tended to mention money and travel (e.g. 'have no ties by the age of 30. Travel round the world. Retire by 45. Be wealthy by 30.')

The major differences in desired life styles were, in fact, between the sexes. Nearly all of the women (90·4 per cent) mentioned marriage among their highlights (some even going into details about the wedding) compared to 75·5 per cent of the men ($\chi^2=9{\cdot}66$, $P<0{\cdot}01$). Children also featured prominently—85·1 per cent of the women mentioned children compared to only 49·7 per cent of the men ($\chi^2=35{\cdot}78$, $P<0{\cdot}001$). But, although marriage and children tended to dominate the highlights list of the female students, this does not mean that career had no place at all. On the contrary, 79 per cent of those who mentioned taking a teaching post indicated their intention of returning to work when their children were old enough.

Significantly fewer male students (27·8 per cent) envisaged careers as school teachers than female students (57·9 per cent) ($\chi^2=24{\cdot}32$, $P<0{\cdot}001$), but they were equally likely to anticipate remaining in education. That is, the men tended to be much more ambitious for headships and posts in higher and further education.

The career ambitions of the male students varied with the type of training they were undertaking—96·6 per cent of the BEd students and 77·6 per cent of the college postgraduates mentioned a career in education compared to 64·7 per cent of the third-year college sample and 48·7 per cent of the UDE group. The latter, however, were more likely to anticipate remaining in schools. Thus, of those indicating a career in education, 89·5 per cent of the university sample and 72·5 per cent of the third-year group envisaged school teaching careers compared to 57·1 per cent of the BEd students and 55·3 per cent of the college postgraduates. The university group tended to mention career rather less than the other groups among their life's highlights.

Discussion

The present study supports the finding of Smithers and Carlisle (1970) that many students in colleges of education would ideally have preferred a career other than teaching. It also suggests that, overall, students in postgraduate training were the more likely to see teaching as their ideal career.

Table 3: Career highlights

	Per cent indicating career peak as:				
Type of course	School teaching	Headship or Deputy Headship	Beyond schools	Outside education	Career not mentioned or intention imprecise
UDE					
m (n=39)	23·1	20·5	5·1	20·5	30·8
f (n=54)	63·0	3·7	nil	7·4	25·9
Colleges—postgraduates					
m (n=49)	20·4	22·4	34·7	6·1	16·3
f (n=31)	59·1	nil	16·1	nil	25·8
Colleges—BEd					
m (n=34)	31·0	24·1	41·4	nil	3·4
f (n=17)	58·3	8·3	8·3	16·7	8·3
Colleges—3rd year					
m (n=29)	41·2	5·9	17·6	20·6	14·7
f (n=12)	41·2	17·6	nil	11·8	8·3

However, there were important differences within the two groups. Among the college students it was the potentially certificated who were disinclined to become teachers. For many of them teaching was not the ideal career, neither were they sure they were actually going to enter the profession. The BEd students, on the other hand, although they did not originally aspire to teaching, seemed to have appreciated that a progressive career in education was open to them. They came out highest of all on the scale measuring strength of intention to teach, and senior posts in education featured prominently among their hoped-for life's highlights.

There were also important differences between students in postgraduate training in the colleges and in the university. Among both groups a high proportion indicated teaching as the ideal career, but, whereas the college postgraduates seemed sure of what they were going to do and looked forward to taking up senior posts in education, the university group were less certain.

This difference may, in part, be related to the thought and effort involved in moving across to a college of education—of the present university sample, 46 per cent were graduates of that university. Also, career did not seem to occupy such a central place in the lives of the university group who tended to give prominence to artistic and sporting achievements. It may be that for these students teaching was seen as the ideal career because of the time it appears to leave free for activities of this kind.

But the similarities between the BEd students and the college postgraduates, and the differences between these two groups and the certificate students and the university postgraduates suggest that anticipated career prospects may be important in commitment to teaching. The BEd students and the college postgraduates will soon have realized that most college of education staff begin their careers in schools and have qualifications not very different from what they, themselves, hope to obtain. College posts will be seen to be attainable. The UDE students, still immersed in the university milieu, may be largely unaware of the range of opportunities open to them. They may also take as their reference group the academics who taught them as undergraduates and who usually make no secret of their poor regard for colleagues in university education departments. The certificate students recognize that they are in a poor competitive position compared to graduates, and it is not surprising that they should be so doubtful about their futures.

Although there were differences in commitment to teaching associated with the different routes of entry, there was also a major difference between the sexes. Consistently the female students were the more committed to teaching and were the more likely to see their careers in education as involving teaching itself. Morton-Williams, Finch and Poll (1966) obtained similar results in their survey of English and Welsh university students.

While it is undoubtedly true that school teaching is a career that fits in well with bringing up a family (about four-fifths of the present sample envisaged eventually returning to teaching), there may also be more fundamental reasons why teaching is more attractive to women. There is abundant evidence that in western society women traditionally come to acquire 'people-orientated' values (Goldsen *et al.*, 1960; D'Andrade, 1967), and teaching is one of the occupations most likely to satisfy these values (Rosenberg, 1957; Cohen, 1969). There is also the point that teaching carries higher prestige and is relatively better paid for a woman than a man (Leggatt, 1970).

What bearing do these results have on the recommendations of the James Committee? At first sight they seem to lend some support for the Committee's preference for consecutive as opposed to concurrent training. The evidence of this study indicates that students in postgraduate training were the more likely to say that for them teaching was the ideal career—presumably because they have had more time and better opportunities of getting to know themselves than the college entrant, and have greater freedom of choice.

However, this is a misleading over-simplification, because ultimately it is the BEd students together with the college postgraduates who seem be the more positive about a career in education.

Commitment to teaching thus has little to with the pattern of training *per se*. It is those who see themselves as being at an advantage who tend to be the most committed in the sense o actually intending to take up teaching. It i doubtful if the James proposals will have muc impact here. Although the Committee proposes kind of all-graduate entry to the profession, teachers will still be differentiated according to the kind of first cycle training which they will have undertaken. It is likely that those who take a degree in the first cycle will be at an advantage and will recognize this during the common second cycle if not before.

The James Report gives relatively little attention to the intrinsic merits of concurrent and consecutive training. Other evidence provided by the present students and reported elsewhere (Smithers and Hellawell, 1972) suggests that both those in concurrent and consecutive training appear to see the concurrent programmes as providing the better professional preparation for teaching.

References

ASSISTANT MASTERS' ASSOCIATION (1970). *The Undergraduate View of Teaching*. London: AMA.

COHEN, L. (1969). 'Identification with a profession: a study of intending teachers', *Educ. Res.*, **12**, 41–5.

D'ANDRADE, R. G. (1967). 'Sex differences and cultural institutions', in MACCOBY, E. E., (Ed.), *The Development of Sex Differences*. Stanford: The University Press.

EVANS, K. M. (1946). 'A Study of Attitudes Towards Teaching as a Career'. Unpublished M.A. Thesis, University of London.

GOLDSEN, R. K., ROSENBERG, M., WILLIAMS, R. M. and SUCHMAN, E. A. (1960). *What College Students Think*. New York: van Nostrand.

JAMES REPORT (1972). *Teacher Education and Training*. London: HMSO.

KELSALL, R. K. and KELSALL, H. M. (1970). 'The status, role and future of teachers', in KING, E. J. (Ed.), *The Teacher and the Needs of Society in Evolution*. Oxford: Pergamon.

KELSALL, R. K., POOLE, A. and KUHN, A. (1970). 'Still a low status profession', *Times Educ. Supp.*, 4. 12.70.

LEGGATT, T. (1970). 'Teaching as a profession', in JACKSON, J. A. (Ed.), *Professions and Professionalization*. Cambridge: University Press.

MANGAN, J. (1970). 'Reluctant students: a statistical survey', *Times Educ. Supp.* (*Scotland*), 30.10.70.

MORTON-WILLIAMS, R., FINCH, S. and POLL, C. (1966). *Undergraduates' Attitudes to School Teaching as a Career*. London: Department of Education and Science.

MUSGROVE, F. (1969). 'Self-concepts and occupational identities', *Univ. Quart.*, **23**, 333–44.

ROSENBERG, M. (1957). *Occupations and Values*. Illinois: The Free Press.

RUDD, W. G. A. and WISEMAN, S. (1962). 'Sources of dissatisfacton among a group of teachers', *Brit. J. Educ. Psychol.*, **32**, 275–91.

SMITHERS, A. G. and CARLISLE, S. (1970). 'Reluctant teachers', *New Society*, No. 388, 391–2.

SMITHERS, A. G. and HELLAWELL, D. E. (1972). 'Teaching: the status split', *New Society*, No. 488, 232–4.

WARNOCK, M. (1970). 'The stigma of schoolteaching', *New Society*, No. 379, 61–2.

VALENTINE, C. W. (1934). 'An enquiry as to reasons for the choice of the teaching profession by university students', *Brit. J. Educ. Psychol.*, **4**, 237–59.

Individuals and Systems

Aspects of Educational Issues in Drug Dependence

David A. Lane, Homewood School, Tenterden, Kent

Summary

There are two important aspects of drugs education which are, unfortunately, often confounded: that is, helping individuals with their problems and education about drugs. This article looks at the first of these and argues that the school can be an important influence both negatively and positively. Both group and individual approaches to the problem are suggested. The second area is to be covered in an article in the next issue of this journal.

The question of drug dependency and the role of the teacher is one that continues to occupy a great deal of attention. Yet in spite of this there still remains much ignorance and a lack of educationally informed research.

In this article I will attempt to deal with the problem of dependency[1] and not with education about drugs, since questions of helping individuals with their problems need to be kept separate from the more general educational issues raised by drugs education. There are, however, a number of useful studies in progress (at the Institute for the Study of Drug Dependence for example) and a review of this aspect has been undertaken by the author to be published in the next issue of this journal (Lane, 1974). As implied in the title, this article deals with two main areas, and this is intended to lend clarity to what are involved arguments.

The individual

The drug addict is someone who is dependent on drugs. True; but unfortunately, it is still the case that lay, and sometimes medical, people lump all the actions of such drugs and all the activities of drug takers into one amorphous mass so that it then becomes possible to talk vaguely about drug addiction as if this were an expression of a simple equation—i.e., drug + patient = addiction. In fact this is not the case (Willis, 1969). The view that drug + patient = addiction is still common and it leads to the view that if you separate drug and individual the problems cease. This makes educational as well as medical nonsense, for we are dealing with personal problems for which some individuals find partial answers in drugs, in the absence of any available perceived alternatives. Willis refers to various studies which show that different people use drugs for a variety of reasons. It must be recognized that not all of these reasons will create a problem. Although, as Mattke and Steinigen (1972) point out, we should not forget the drug itself and its actions.

Why exactly some problems lead to dependency is not clear and the whole area falls into that unfortunate category of multi-causal. The label multi-causal is unfortunate because it makes it difficult to unravel the varied contributory factors. However, something that becomes apparent from the statement by Willis is that mass media slogans which lay the blame completely on drugs as if in some magical way they transformed secure well-adjusted individuals into babbling idiots, are misleading at the least, and, at worst, are positively dangerous. It is sad that the reported statements of some educators at conferences are full of such scare-mongering. It must be recognized that a certain amount of drug use takes place. The occasional user should not be the subject of the hysterical overconcern to which he is at present subjected.

Acknowledgement
The author would like to thank Dr. Anthony Wood of the Health Education Council for his advice in the preparation of this article.

---

[1] That is those aspects of dependency which manifest themselves in terms relevant to education in the school system, but not dependency in its entirety. I am not concerned here with the occasional user. The importance of the term 'dependency' in this article is to the extent that an individual's persistent use of a substance or substances is such that the organization of his overall behaviour is disordered by this usage.

If there is any agreement in the field of dependency it is that the drug-dependent individual in some way has an inadequate personality. Laurie (1969) came to this conclusion in his review of the evidence and Dale Beckett in a pamphlet for parents published for the Association for the Prevention of Addiction (1972), stresses the fact that a secure individual is unlikely to become dependent on drugs. Similarly, from the research evidence it is becoming obvious that some sort of personality problem is at the basis of such drug use. A useful summary of the evidence is available from the Teachers Council on Drugs Education. The existence of problems such as family background and poor education is clearly indicated in this summary although those of good education, etc. are also drawn into dependency (Hicks, 1972). However, the difficulty in isolating the actual personality variable which might lead to dependency was made clear in an article by Teasdale, Seagraves and Zacune (1971) in which they pointed out that although as a group the drug dependent score highly on Eysenck's Psychotism Scale only a minority of drug users actually answer the questions in the psychotism direction. Other data show, again as a group, that drug users score highly on scales or neuroticism (Halstead and Neal, 1968; Rosenberg, 1969).

This difficulty is perhaps to be expected since several of the factors which may lead an individual to try drugs are symptoms of normal adolescent experimentation (Lane, 1970; Richter, 1970) and this is, in part, a popular view (Craven and Birdwood, 1971). But, the idea that individuals may be led to dependency rather than minor experimentation as part of normal adolescence is as incomplete an answer as is sole reliance on personality factors as the key. Although many people agree about this, they aften do not realize why this is so or they fail to accept the implications of this. If any of the above factors are to result in drug dependence the individual must first have access to drugs and learn attitudes favourable to their use and discover in using them that they work and are to be valued.

Even if individuals do have appropriate access this does not mean that they will become dependent. Some individuals experiment with drugs, others do not, and some become dependent while others simply use drugs as part of a culture and cannot be viewed as disordered individuals simply because of this use. It is only if individuals learn to place a sufficiently high value on drug use, and/or culture, so that drug use becomes a powerful motivating factor in their behaviour that there is a danger that it will become the major motivational factor available to them. If it were a question simply of normal experimentation leading to disorder because of the evil and automatic effect of drug use on the personality of the user, the figures for dependency would be astronomical. Unfortunately, various myths accompany drug use of which this is one. This particular myth leads to the view that this is purely a medical problem, as Tony Ryle points out (1972). The addict was a product of family and societal problems which went back for years and it is not helpful to let society off the hook by saying that it is a medical problem.

The cultural aspects, i.e. access, attitude and motive, of dependency makes available to the educator the general area of the sociology of deviance as a theoretical base which may throw light on the learning of drug-dependent behaviour. Sutherland and Cressey (1955) maintain that individuals learn deviant attitudes in intimate personal relationships, thus we may look to such relationships for the primary learning of dependent behaviour rather than from impersonal sources such as television. Merton's theory of anomie (1949) and, more particularly, the reformulation of it by Cloward (1959) to include the idea of differential access, emphasizes that individuals adapt to the situation they are in, choosing courses of action from those available. For those who fail within the social system, access to the different options is often limited. Withdrawal from society is one form of adaptation mentioned by Merton which is relevant to dependency, and this is an option which has a tradition of over a thousand years. But it must be remembered that patterns of drugs use change very rapidly and the information which is most useful is that related to the area in which the teacher works. There is not a typical drug-taking group the stereotype of which can be applied universally. O'Sullivan (1972) for example, has outlined the different groups that can exist in an area of one square mile of London.

Interim conclusion

In terms of the individual, it is simply not adequate to think of the drug-dependent individual as someone whose problem is drugs.

Rather, he is someone who is peculiarly susceptible but not necessarily predisposed to dependency because he finds it difficult to cope with stress, and, who, because of involvement with a group or individual who has favourable attitudes to drug use then learns to use drugs successfully at least initially, to obliterate or overcome his problems or to find answers to whatever it is he is looking for.

The school as a social system may, in part, force or usher the individual to select the option of dependency by denying the individual ease of access to alternative roles.

Systems

We have seen that the development of dependency requires appropriate access, attitudes and some value derived from drug use not as readily supplied from some other source. The process for the development of the potential for such patterns of dependency, may start long before drugs are ever used. Although this article concentrates on influences within the school, this does not mean that there are no other contributory influences such as the family, youth groups or mythology. In this section the extent to which schools can bring about a situation in which all of the above can occur is considered. This does not mean that all schools stand condemned as the causal agents in drug dependency but rather that there are factors in the school system which, given appropriate circumstances, contribute to the creation of the conditions for dependency.

All organizations are constructed to achieve particular goals, schools are no exception. The achievement of these goals is dependent on a number of factors including informal as well as formal influences. Therefore the influence of informal latent cultures within a school are as much a part of achievement within the school as are formal policy decisions. These influences are particularly important when there is conflict, as there always is to a greater or lesser extent, between personal needs and the position the individual holds within an organization, be he teacher or pupil. The position an individual holds (his roles) imposes certain demands and obligations but the means for him to fulfil these organizational demands may not be easily available. Thus, the teacher may find conflict in his concern for an individual with problems as against his position as a member of the staff. The pupil may find himself unable, in ways approved by the school, to meet the demands for success. 'Role dilemmas have their sources both in organizational structure and in individual personality. Similarly, both structure and personality influence the varied forms of adoption that are achieved' (Levinson, 1969).

Individuals bring their own expectations and values with them to school and find within the school informal and formal influences. There is thereby an ever present potential for conflict for staff and pupil where such varied interpretations exist, alongside role dilemmas. The existence of such conflict lessens the potential for the realization of the formal goals of the organization and increases the potential for the realization of alternative informal goals.

The individual in the school interprets the behaviour of others and consequently his action is subject to change. Most teachers have probably been made aware of the interpretive nature of their students' behaviour at one time or another. A situation exists in which individuals are not simply doing as they are told but are taking courses of action from within a system of options based in the relative advantages of the behaviour to themselves. Just think of what pep-pills may appear to offer to the individual anxious for immediate gratification when his confidence is shaken by failure (or perhaps even the young sportsman desperate for success). The existence of this interpretive learning process has been demonstrated in various studies of which Becker's study (1963) of drug use is the most relevant to the present discussion.

Bearing in mind what was said of the conditions necessary for the development of dependency in the first section of this article, consider how favourable the conditions might be for the development of latent cultures when individuals fail to achieve success in ways and at a level they consider appropriate within the school. In fact, anybody who considers himself a failure, whatever the reality of the situation, must be considered at risk given appropriate circumstances. In bringing together individuals with similar problems and values and of roughly similar age, the likelihood of a group based on deviant and possibly drug-based culture being formed is increased.

In spite of the popularity of the image of the pusher it is a fact that most people are introduced to drug use by friends. In a study of Heroin users Willis (1969) found that this was true in nearly every case. Even without a coherent group, a quasi-group relying on the mythology of drugs and some personal contact might be formed, and which is capable of resisting pressures to conform to formal demands because of the levels of ambivalence and apathy which exist in our schools. Individuals thus find within such groups the success denied elsewhere (Mays, 1972; Lane, 1973).

We have, then, a situation in which not only does the school structure create conflict through pressures of achievement, etc., which individuals are unable to meet, but it provides the means for producing contradictory alternatives to its goals. However, the availability of drugs within schools is far less than is popularly supposed, probably no more than 10 per cent have some knowledge of drug use and a larger section of this group have only a passing acquaintance (Wiener, 1969). This together with the economics of the matter (raids on mothers' librium in the first-aid cabinet, however, seems to be a growing problem) is a limiting factor. But, the basis for later drug use is laid in many cases at school even though drugs themselves may not be used.

Possibilities in the need for educational action

If it is true that the school can be a factor leading to dependency it is also true following the same line of argument that the school can play an important part in preventing the development of problems leading to dependency.

Elsewhere (1970) I have outlined the role the teacher can play by providing alternative successes and argued that the professional teacher should be capable of recognizing the symptoms of failure to cope with stress situations which the adolescent faces. This same article provides a list of such indications. The techniques suggested, such as using a person's interest in something to get him integrated into a non-deviant group, are derived from social group work theory as applied to youth and community work (MacCullough and Ely, 1969; Matthews, 1966; Morse, 1968). This has certain advantages over individual counselling, although it is usually used in association with it, in that it does not rely as heavily on the form of intimate social relationships which are difficult to establish under the pressures of the school system and it also allows larger numbers to be helped indirectly. The whole structure of individual counselling within schools is so confused at the moment that little of a constructive nature can be said.

The most useful approach for the present is united action jointly with the health and social service agencies particularly the local Medical Officer of Health as a first step, preferably through the work of the Drug Liaison Committees where they exist. However, recognition of the responsibility of the teacher for those pupils who fall foul of the system for one reason or another is the start that is needed. Rather than identifying so-called drug users, teachers should identify those who need help for whatever reason. Don't go round looking for addicts, look for individuals who need help. The implementation of the Children and Young Persons Act 1969 may lead to greater flexibility in treatment for the pupil who comes up against the law but again matters vary so much from one area to another. Find out what is going on in your area before you offer pupils to the mercies of the local courts, that is the only advice that can be given. For once 'labelled' they may be driven to further extremes (Becker, 1963).

However, it is not simply a question of individuals with problems needing counselling. For while most people would recognize the right of the individual having difficulties with his GCE studies to sympathetic counselling, many would deny the same right to the individual caught taking drugs regarding this as a matter for repressive discipline. Drugs do involve a number of other factors which are educationally relevant and must be faced. The ambivalent position of the school in relation to the requirements of the wider community as against its commitments to the individual is an example of this.

If teachers constantly grass on pupils can they wonder if no one will confide in them. This is, of course, not the result of bloody-mindedness on the part of the educators but simply a further example of role-dilemma. Teachers do have a responsibility legal and moral to all their pupils not only those with problems. They also have a responsibility to the wider community, and their employing authority. A local education authority is hardly going to look favourably on a teacher who fails to report the existence of a drugs ring because he wants to retain the participants' confidence!

If the teacher is to provide education, that is to provide the individual with a basis from which he can build for himself, he has to accept the implications of that decision, i.e., that the individual must be free to choose. If the individual is to become independent of the teacher his choice must be his own, and not automatically that of the teacher. But this level of choice is not available in many situations. The school which allows individual deviance will provide a situation favourable to such an approach. In the school with restrictive aims and the power to select its intake according to its own demands such an approach may be neither desirable nor useful. Such a school because of the greater concensus between parents, staff and pupils may more easily indoctrinate negative attitudes to drug use—something the permissive school could never do. This does not imply a value judgement on the relative merits of each but simply an acceptance of the fact that different educational situations give rise to the need for different approaches.

The problem of dependency does not arise over night and there are strict limitations to the action any individual teacher can take. But as part of a co-ordinated policy in conjunction with other local authority services, the school, providing the factors mentioned in this article are taken into account, can greatly reduce the problems of dependency. For those not dealt with here, e.g., the totally inadequate, the hardened sociopath (Yablonsky, 1962) the school can offer little help but through partnership with the other services lies the best hope for these groups as well.

There is no solution to the problem, not because there is nothing that can be done, but because there is no one correct course of action. Each school must as a matter of policy find its own solution based on knowledge of the local situation. Since dependency is the product of a variety of factors, of which school structure is one, then that structure must be taken into account when planning action. It is a responsibility that schools cannot afford to shirk for even if they take no action they are still a part of the development of patterns of dependency and the individual needs the help of the school to overcome it. What could be more of an educational issue than that the school is part of this problem and what more need could there be for educational action than that the school contributes to the development of delinquent behaviour? (Power, *et al.* 1967; Mays, 1972). The school *is* part of this problem and the continuing lack of action and denial by many education authorities of the problem is a sad reflection on the professional awareness of educators. Authorities, schools, colleges, and universities have yet to take positive action in this area, both in terms of education and counselling, although there are indications that concern and interest is giving way to positive action (Report of the national survey of an education working party of the Hospital Centre, 1973).

Fortunately, ignorance is giving way to action but the success of the odd school or education authority is not enough, for until schools accept responsibility for the deviant he will be driven to further extremes until recovery becomes beyond reach. If we are to deal with failure, brilliant high spots will not do. An achievement-based society valuing success through education will create frustration for the rejected. If we do not believe in them they will retaliate or withdraw (Lane, 1970).

References

BECKER, H. S. (1963). *Outsiders: Studies in the Sociology of Deviance.* New York: The Free Press.

BECKETT, D. (1972). *Beat the Drug Plug—Advice to Parents.* London: The Association for the Prevention of Addiction.

CLOWARD, R. A. (1959). 'Illegitimate means, anomie, and deviant behaviour', *Am. Soc. Rev.*, **24**, 2.

CRAVEN, M. and BIRDWOOD, G. (1971). *Drugs in Schools.* London: The Association for the Prevention of Addiction.

DRUG DEPENDENCY DISCUSSION GROUP (1973). Report of the Education Working Party: The Hospital Centre.

HALSTEAD, H. and NEAL, C. D. (1968). 'Intelligence and personality in drug addicts; a pilot study', *Br. J. Addict.*, **63**, 237.

HICKS, R. C. (1972). In *Drugs.* Manchester: Teachers Council on Drug Education.

LANE, D. A. (1970). 'Drugs: the role of the teacher and youth leader', *Community Health*, **1**, 6, 327.

LANE, D. A. (1973). 'Drug abuse: deviance, mental illness or what?', DDDG Reprint. London: The Hospital Centre.

LANE, D. A. (1974). 'Drugs education: questions and answers', *Educ. Res.*, **16**, 2, to be published.

LAURIE, P. (1969). *Drugs.* Harmondsworh: Pelican Books.

LEVINSON, J. D. (1969). 'Role, personality and social structure', in COSER, C. A. and ROSENBERG, B. *Sociological Theory.* London: Macmillan.

MacCullough, M. K. and Ely, P. J. (1969). *Social Work with Groups*. London: Routledge & Kegan Paul.

Matthews, J. (1966). *Working With Youth Groups*. London: University of London Press.

Mattke, D. J. and Steinigen, M. (1972). 'An analysis of steet drugs', *Drugs and Society*, **2**, 2, 17.

Mays, J. B. (1972). *Juvenile Delinquency, The Family, and The Social Group*. London: Longman.

Merton, R. K. (1949). *Social Theory and Social Structure*. New York: The Free Press.

Morse, M. (1968). *The Unattached*. Harmondsworth: Penguin.

O'sullivan, P. (1972). 'A square mile of drugs use', *Drugs and Society*, **2**, 2, 12.

Power, M. J., Alderson, M. R., Phillipson, C. M., Shoenburg, E., Morris, J. N. (1967). 'Delinquent schools', in *Crime, Deviance and Social Sickness*, a New Society Social Studies Reader. London: New Society Publications.

Richter, D. (1970). *Drug Problems and their Management*. London: The Association for the Prevention of Addiction.

Rosenberg, C. M. (1969). 'Young drug addicts: background and personality', *J. Nerv. Ment. Dis.*, **148**, 65.

Ryle, R. (1972). 'Drug mythologies', Reprint 728. London: The Hospital Centre.

Sutherland, E. H. and Cressey, D. R. (1955). *Principles of Criminology*. Philadelphia: J. B. Lippincott & Co.

Teachers Council on Drug Education (1972). *Drugs*: Manchester.

Teasdale, J. D., Seagrave, R. and Zacune, J. (1971). 'Psychotism in drug users', *Brit. J. Soc. Clin. Psychol.*, **10**, 2, 160.

Weiner, R. S. P. (1969). *Drugs and School Children*. London: Longman.

Willis, J. H. (1969). 'Drug taking—myth or reality?', *Community Health*, **1**, 3, 121.

Yablonsky, L. (1962). *The Violent Gang*. Harmondsworth: Penguin.

Useful addresses

The Association for the Prevention of Addiction, 33, Long Acre, London, W.C.2.

Teachers' Council on Drug Education, 437, Royal Exchange, Manchester.

The Institute for the Study of Drug Dependence, 3, Blackburn Road, London, N.W.6.

The Drug Dependency Discussion Group at the Hospital Centre, 24, Nutford Place, London, W.1.

Structure and Process in Secondary Schools

The Academic Impact of Educational Climates

Edward L. McDill and Leo C. Rigsby

Although educators have long believed factors such as class size, school finances, and quality and quantity of physical facilities are valid indicators of school effectiveness, the authors find that this is not the case. Based on a thorough study of the educational and social climates in representative high schools throughout America, the authors disprove a number of widely held theories as to what factors do influence student performance.

They demonstrate a direct relationship between the quality of high schools and the academic performance and aspirations of students. In addition, they show that this relationship is not tied to a student's personal, familial, or socioeconomic characteristics. Most of the authors' analysis is conducted separately by sex, and this factor produces a number of important findings about the differential impacts of school, peer group, and family on the academic productivity of girls and boys.

This innovative study has significant implications for educational policy, and should serve as a stimulus to new thinking among educators and administrators alike.

Edward McDill is Professor and Chairman of the Department of Social Relations at The John Hopkins University. Leo C. Rigsby is Assistant Professor of Sociology at Vanderbilt University.

244 pages £4.05

The Johns Hopkins
University Press
2-4 Brook Street
London W1

The Three-minute Reasoning Test

a Re-evaluation

James Hartley, Department of Psychology, University of Keele
Frank W. Hogarth, College of Education, Madeley, Staffs
Richard L. Mills, Fogo School, Fogo Island, Newfoundland

Summary
In a previous paper (Hartley and Holt, 1971) the development of a three-minute reasoning test was described, and its validity and reliability assessed. In this paper three experiments are reported which look in greater detail at these aspects of the test in different situations. The results of the experiments indicated that the test was less valid and was less reliable than was originally anticipated.

In an earlier paper, Hartley and Holt (1971) described their attempts to develop Baddeley's (1968) Three-minute Reasoning Test in order to make it more suitable for work with school children. Such a test, it was argued, would be of use in situations where there was insufficient time to administer a full-length test of intelligence, such as the AH4 (Heim, 1967), or where the matching of groups had to be based on data which were out of date (e.g., previously obtained 11+ measures). Hartley and Holt found a significant correlation of +0·70 between scores on their simplified version of the Three-minute Reasoning Test and the AH4 (N=164), and they also found that correlations between various performance measures and scores on the AH4 were of a similar size to those found between the same performance measures and scores on the Three-minute Test.

However, a test-retest correlation of +0·66 (N=28), obtained over a seven-day interval, indicated that the reliability of the Three-minute Test might be low. In their discussion of results, Hartley and Holt commented upon a practice effect, especially for lower ability pupils, and suggested that the test might reasonably be extended to four minutes, and then scored (say) from item number 17 (i.e., from page 2 of the test). Baddeley (personal communication), however, commented that scores on the earlier part of the test might be more crucial for predicting performance than later test scores.

In this paper three experiments are described which examine some of these questions. Experiment One compares the results obtained from the Three-minute Test with those obtained when the test was extended to four minutes, and the test scored from item 17. Experiment Two compares the predictability of performance from scores obtained from the first minute with that obtained from the total Three-minute Test. Experiment Three describes the results obtained with a further simplification of the Three-minute Reasoning Test. Some more general comments follow the presentation of these results from the three experiments.

Experiment One

In the earlier paper, Hartley and Holt (1971) described an experiment (Experiment Three) which was basically concerned with programmed instruction, but in which the investigators also collected data from the Three-minute Reasoning Test. The results of the experiment are here reported as the first part of Table 1. It can be seen that in that experiment, as the investigators pointed out, the correlations between the Three-minute Reasoning Test and other measures of achievement (11+ data, exam data and performance on the programme) were variable and inconsistent, particularly in relation to the 11+

measures, although the Three-minute retest data appeared to be superior predictors of pr-ogramme performance.

Experiment One in this paper was a replication of this experiment, except that in this study one group of pupils chose their own partners to work with through the programme, and in a second group partners were formed by the teacher. (Since there was no significant difference between the results obtained from the two groups —see Hogarth and Hartley (1973)—the data have been pooled in the investigation to be reported here.) In addition, in this study, a change was made in that the time allowed for the Three-minute Reasoning Test was extended to four minutes, and pupils indicated, by drawing a line across the page, which point they had reached after the third minute. The test was administered before and after the experiment, and the test-retest interval was four to five weeks—unlike the first study where it was approximately seven days. The second half of Table 1 shows the results obtained from this investigation.

Four main points emerge from Table 1: (i) the correlations from the Three-minute Test were more consistent than were those obtained in the 1971 experiment; (ii) the retest data produced higher correlations than did the test data; (iii) the scores obtained in the Four-minute Retest, scored from item number 17, produced the highest correlations (for the reasoning test) but, (iv) in this study the 11+ data produced the highest correlations of all. A fifth observation was that the test-retest correlation found for the Three-minute Test was 0·32 ($p<\cdot 05$), and for the Four-minute version it was 0·38 ($p<\cdot 01$): both of these correlations are substantially lower than that of 0·66 reported in the previous paper.

Finally, we may note in connection with this experiment, that two other measures of reliability were calculated in addition to the test-retest one just reported. In the previous paper, Hartley and Holt reported an odd-even split-half correlation of 0·94 for the test, indicating a high internal consistency. Start (personal communication) suggested that a median split-half correlation would be a more appropriate measure to use for such a test. In this study we calculated both odd-even split-half and median split-half correlations for the test (+0·91 and +0·91 respectively) and for the retest data (+0·85 and 0·93 respectively). Thus both these measures support the previous finding of a high internal consistency for the test.

Experiment Two

In this study the Three-minute Reasoning Test was administered to 61 male university students at Memorial University, Newfoundland, both before and after an experiment on reading and typographical layout. These students were aged mainly between 16 and 18 and had a wider ability range than is customary in the UK. Two performance measures, reading speed and comprehension were made in this experiment, which lasted for about twenty minutes. The students were tested singly, or in groups of twos and threes,

Table 1: Intercorrelations obtained between the Three-minute Reasoning Test and other measures of performance

(a) *The 1971 Study*	11+data	Total Exam. Scores	Chemistry Exam. Score	Programme Test Score/25	Programme Retest Score/25
11+ data (N=45)	—	0·36*	0·55**	0·28*	0·33*
3 Min. Test (N=45)	0·36*	0·23	0·34*	0·08	0·32*
3 Min. Retest (N=28)	0·45*	0·15	0·07	0·50**	0·49**
(b) *The 1972 Study* (N=57)					
11+ data	—	0·65**	0·49**	0·59**	0·52**
3 Min. Test	0·29*	0·32*	0·21	0·20	0·17
3 Min. Retest	0·42**	0·36**	0·35*	0·36**	0·26*
17–4 Min. Test	0·30*	0·28*	0·22	0·19	0·11
17–4 Min. Retest	0·46**	0·44**	0·38**	0·37**	0·29*

*=$p<\cdot 05$.
**=$p<\cdot 01$.

Table 2: Test scores obtained for each one-minute section of the Three-minute Test, test-retest correlations, and internal consistency measures (N=61).

	Scores 1st attempt		*Scores 2nd attempt*		*Test-Retest correlations*
	X̄	s d	X̄	s d	
1st Minute	11·11	4·59	13·07	4·59	0·73
2nd Minute	9·74	3·29	11·62	3·69	0·62
3rd Minute	9·72	3·94	11·46	3·80	0·62
Total	30·57	9·92	36·15	10·67	0·79
Odd–even split correlation	0·93		0·90		
Median split correlation	0·92		0·92		

and the instructions (basically the same as those reported by Hartley and Holt, 1971), were tape-recorded and played back during each session. The instructions were modified slightly in that the students were requested to circle the number of the item on the test they had reached when they heard a click from the tape-recorder, with the instructions 'One minute, Circle', 'Two minutes, Circle', etc.

Thus in this study scores for each minute, total scores, and retest scores were available. Table 2 summarizes these data.

Three measures of reliability are reported in Table 2—a test-retest, an odd-even split-half, and a median split-half. These results indicate again the test has a high internal consistency, but that again the reliability of the test over time (this time twenty minutes) is not high.

It is interesting to observe here, in connection with Baddeley's comment, that scores in the first minute are higher than those obtained in the second and third minute for both sets of test results. An analysis of the number of errors made did not indicate that they significantly increased in the second and third minute sections of the text: indeed, the error rate was constant.

Table 3 summarizes the intercorrelations between the scores obtained on sections of the test and the performance measures of reading speed and comprehension for both the first and second attempt.

The data in this table suggest that with the first attempt at the test the scores obtained in the first minute were just as good predictors as were the total scores, but that for the second attempt, the total scores were slightly better. In general the correlations with the second attempt at the reasoning test were higher than those obtained with the first attempt (only 3 correlations out of the 14 reported are lower).

Experiment Three

Whilst watching the Three-minute Test in use it was observed that part of the difficulty of the test seemed to derive from its layout rather than from its content. Testees are required to judge whether a sentence (e.g. A comes before B) describes correctly or not the order of a letter pair (e.g. BA), and to mark their conclusion with a tick in a True or False column. It seemed that part of the difficulty lay not in deciding whether the sentence was correct or not, but in putting the tick in the appropriate column. To test whether this difficulty was only one of layout, a revised form of the Three-minute Test was constructed in which the testees were required simply to mark each description of the letter pairs right or wrong—using a tick or a cross. If the layout of the initial test caused difficulty, it would be expected that higher scores would be obtained on this simpler version: if, however, the difficulty was also related to verbal reasoning it would also be expected that correlations with a verbal reasoning test would be lower with the

Table 3: The correlations between scores on sections of the Three-minute Reasoning Test, reading speed, and comprehension. (Correlations for the second attempt are shown in brackets.)

	Comprehension	1st Min.	2nd Min.	3rd Min.	TOTAL
Reading speed	0·23	·24 (·24)	·08 (·17)	·25 (·24)	·24 (·27)
Comprehension		·16 (·15)	·22 (·35)	·14 (·20)	·16 (·20)
1st Min.			·58 (·65)	·54 (·73)	·84 (·87)
2nd Min.				·70 (·63)	·82 (·86)
3rd Min.					·86 (·87)

Correlations to be significant at $p < \cdot 05$ must exceed 0·21 when N=61.

Correlations to be significant at $p < \cdot 01$ must exceed 0·30 when N=61.

simpler version. Experiment Three tested these assumptions.

Four mixed-ability second-year forms in a boys grammar school were used as subjects in this experiment. The two versions of the Three-minute Test were administered at random in each form, and were then completed following appropriate instructions as regarded marking the answers. Then each boy completed the first half of the AH4 intelligence test (Heim, 1967). This part of the test concentrates on verbal and numerical problems. (This test was also used in the previous investigations reported by Hartley and Holt, so this study can be regarded as a further replication of part of the original investigation).

The data for each form is presented in Table 4 in order that some of the inconsistencies may be observed. However, speaking generally, it may be observed that the simplification of the initial version of the Three-minute Test did result in higher scores ($t=2{\cdot}55$ $p<{\cdot}01$, one tail test), but that the correlations were much the same order of size. The correlation of 0·48 between the AH4 and the standard version of the Three-minute Test is again lower than that reported (0·73) in the earlier paper of Hartley and Holt.

Concluding remarks

The three experiments outlined in this paper taken as a whole present a picture which is far less encouraging than that originally envisaged by Hartley and Holt. In brief, it appears that the Three-minute Test, although internally consistent, is not very reliable, even over short periods of time. It is clear (e.g. see Table 2) that practice effects can occur, and that (see Tables 1 and 3) better results are obtained with scores from the second version rather than from the first version of the test. These results, which may perhaps partly result from a reduction in anxiety in the test situation, unfortunately make the prospect of finding a short but reliable test, quick to score and easy to administer, somewhat further off. In addition, the results of Experiments One and Three indicate that there are likely to be considerable fluctuations in terms of correlations with the AH4 or other intelligence tests. Indeed, the correlations reported in these experiments are much lower than those initially reported by Hartley and Holt. It appears that situational and personality factors influence the results—some testees palpably guessed the answers, others failed to treat the test seriously—

Table 4: The mean scores and the intercorrelations obtained from the AH4 and the two versions of the Three-minute Reasoning Test

		AH4	*Initial Three-minute Test*	*Correlation*	*AH4*	*Revised Three-minute Test*	*Correlation*
	$\bar{x}$	37·0	31·7		31·7	28·8	
Form 1	sd	7·2	10·0	0·68	6·4	9·0	0·61
	N	11	11		10	10	
	$\bar{x}$	35·4	28·7		34·6	33·5	
Form 2	sd	5·7	6·1	0·34	4·4	6·9	0·33
	N	12	11		11	11	
	$\bar{x}$	36·5	25·7		34·9	31·6	
Form 3	sd	3·5	10·3	0·49	5·7	9·5	0·17
	N	11	10		10	10	
	$\bar{x}$	35·6	25·4		37·0	36·8	
Form 4	sd	6·3	8·2	0·42	4·8	7·9	0·52
	N	11	10		10	10	
	$\bar{x}$	36·0	27·9		34·5	32·9	
TOTAL	sd	5·9	9·1	0·48**	5·7	9·1	0·45**
	N	45	45		41	41	

**$p<{\cdot}01$.

and there were discipline problems. Putting it briefly, however, in its present form the Three-minute Reasoning Test has turned out to be a less reliable and a less valid instrument than was first anticipated.

References

BADDELEY, A. D. (1968). 'A three-minute reasoning test based on grammatical transformation', *Psychonomic Science*, **10**, 10, 341–2.

HARTLEY, J. and HOLT, J. (1971). 'The validity of a simplified version of Baddeley's Three-minute Reasoning Test', *Educ. Res.*, **14**, 1, 70–3.

HEIM, A. H. (1967). *AH4 Group Test of Intelligence* (Revised Edition). Slough: NFER.

HOGARTH, F. W. and HARTLEY, J. (1973). 'Programmed learning: mixed ability versus own-choice pairs', *Programmed Learning and Educ. Technol.*, **10**, 2, 69–76.

The Maturation of Spatial Ability in Geography

A. J. Heamon, Gipsy Hill College, Kingston upon Thames

Summary

The study reported here attempted to discover by employing different geographical techniques at which stages certain abilities—generalization and comparability—develop in children. Eighty children aged between 8 and 14 were tested in several ways using maps, pictures and photographs to see if they could (a) generalize from a distribution and (b) compare two separate distributions. In both cases tests revealed an association of performance with age: by the end of the primary school about 50 per cent of the children could perceive a relationship between two generalized distributions, and the test to compare two separate distributions, above 'O' level standard, was well executed by 25 per cent of the 10-year-olds.

Introduction

Geography has always been remarkable for its diversity of content, and whether this is a matter of notability or of notoriety has been the subject of endless discussion. As far as children's thinking is concerned, however, there is reasonable agreement that the discipline provides valuable experience in the handling of spatial distributions including maps, photographs, models and the 'out of doors'. Such experience includes, among others, the following intellectual processes:

1. The generalization of a distribution of discrete items, e.g. seeing a number of spring-line villages as a band of settlement.
2. The comparison, with the possibility of correlation in mind, between two distributions; either on the ground, on a single map, or represented for the purpose of analysis on different maps.

The general purpose of the study described below was to discover the stages by which the performance of these activities develops in children.

Eighty children of Abingdon and district were tested individually. The total comprised four groups each of 20 children aged 8, 10, 12 and 14 years. The children were all born in September and it was possible to test nearly every child with the right age and birthday attending maintained schools in the area. The method of sampling was intended to ensure a representative distribution of ability without the necessity for pre-judgement which would have arisen had a selection been made on the basis of intelligence quotients.

Generalizing a distribution

Three tests of the ability to generalize were used.

A. The children were given a colour photograph of a town with mountains behind (Keswick). They were then asked to draw one shape to show the outline of the town on a piece of paper the same size and shape as the photograph.

B. A large 25 inch map measuring approximately 38 inches × 25 inches with all the buildings of a village near the centre coloured red was shown to the children. They were then given a piece of paper with the shape of the map frame already drawn and asked to draw one shape to show the outline of the village.

C. The material for the third test consisted of a 6 inch map measuring approximately 23 inches × 19 inches with several villages, the houses of which were coloured red, arranged in a discontinuous belt running from W.N.W. to E.S.E. Rivers were coloured blue, ground over 200 feet brown and woods were shown by green dots. The children were given a piece of paper with the shape of the map frame already drawn and asked to draw one shape to show where most of the houses were.

The greatest range of performance occurred on Test B and this was used as the basis of the marking scheme. It was considered that six categories of performance could be recognized:

1. Unable to grasp the idea of representing the data by a single outline.

2. Some idea of what was required but no enclosed shape formed. Typically, children whose work was in this category drew the outside edges of individual houses and linked some together but eventually began to draw the houses separately.

3. Children whose work was in this category drew an enclosed shape but missed the structure of the village which is an inverted and lop-sided 'T'.

4. Work in this category shows that the children had formed a mental image of the shape of the village but relied on the use of printed field boundaries and other lines in order to draw it. The drawings have numerous sharp angles and exaggerated details.

5. Only a few lines used for guidance in drawing.

6. A smoothed outline apparently based on a firm percept of the village as a whole.

These categories, grouped as 1—3, 4, and 5—6 were also applied to the work done on Tests A and C. A frequency table of the results is given in Table 1.

Table 1 can be used to distinguish between the ability to form a percept of a generalized outline (Categories 4, 5 and 6) and the ability required to reproduce it accurately (Categories 5 and 6). The latter is presumably dependent on short term visual memory and is clearly much later in development than the former. The implication for teaching is that, whilst most children from the third or fourth year of the primary school onwards can be given the opportunity to make generalizations based on spatial data, the process should be carried out through oral discussions for about two years.

Table 2: Difficulty of tests compared with size of area covered

Tests in order of difficulty	*Area over which the data are distributed in sq. cms.*
C	20
A	32
B	420

The marked differences in attainment on the tests also deserve notice. For most of the children Test B was the hardest of the three and Test A was harder than Test C. Although other factors no doubt had some influence it looks as though the size of the area over which the data are distributed is of over-riding importance in this instance.

These results suggest that when children reach the stage for written and graphic work involving spatial generalization they should begin by using maps and pictures smaller than, say, quarto size.

Comparing two distributions

The following tests were designed:

D. The children were shown a colour photograph of a large hill without vegetation (Ayers Rock, Australia) but with a line of trees near the base. They were then asked to pretend that the experimenter could not see the trees, and to describe the position of them. Supplementary questions were asked when necessary.

Table 1: Test of ability to generalize: categories of performance by age

Age	Test A *Categories*			Test B *Categories*			Test C *Categories*		
	1–3	4	5–6	1–3	4	5–6	1–3	4	5–6
8	8	6	6	11	7	2	8	7	5
10	6	5	9	4	11	5	1	9	10
12	4	2	13	6	8	6	3	10	7
14	1	4	15	2	8	10	0	4	16
Total	19	17	43	23	34	23	12	30	38
		79			80			80	

All three tests show an association of performance with age which, using the X^2 test, exceeds the ·05 level of significance for tests A and B and the ·001 level of significance for Test C.

E. Using the map already described in connection with Test C, the children were asked what they noticed about the position of nearly all the woods.

It will be noted that in Tests D and E the distributions were superimposed. The material described below was designed to test the ability to correlate separated distributions.

F. A number of irregular shapes were drawn on a piece of ¼ inch squared paper measuring 5¾ inches × 3¾ inches, i.e. with 345 squares. The general appearance was that of a mainland and archipelago which covered exactly 50 of the squares. A series of four dot distributions was then prepared, each distribution consisting of 50 dots on a piece of squared paper the same size as that given above. The first used 40 of the squares in the areas covered by the shapes, the second 30, the third 20 and the fourth 10, giving overlaps of 80 per cent, 60 per cent, 40 per cent and 20 per cent respectively. All the distributions were then transferred to plain white paper and mounted on a grey background. A similar set was prepared with a dot distribution as standard and four shapes in descending order of correlation. A third standard consisting of a black line network was also drawn for comparison with dot distributions. In each case the children were asked to pick out the pattern of shapes or dots which showed most overlap, the one which overlapped least and the one which was second best.

The scores on Tests D, E and F are not directly comparable. On Test D there were three main categories of answer:

1. No evidence that a relationship was seen between the trees and any other feature in the picture.

2. The trees were seen as related either to the hill or to the flat ground.

3. The trees were related both to the flat ground and to the hill.

Table 3: Type of answer on comparability question by age

Age	Categories 1	2	3
8	11	7	2
10	7	9	4
12	5	13	2
14	0	14	6
TOTAL	23	43	14
		80	

There is an association of performance with age which exceeds the ·01 level of significance on the X^2 test.

The frequency distribution by age is shown in Table 3.

Test E called for perception of the relationship of woods to high ground and the answers are classified simply as 'not seen' or 'seen'.

Table 4: Perception of a relationship by age

Age	Not seen	Seen
8	17	2
10	11	9
12	10	10
14	2	18
TOTAL	40	39
	79	

The association of marks with age is significant at the 0·01 level.

Test F carried a possible maximum of 9 marks and because there were only four possibilities in each set a score of between 2 and 3 could be obtained by chance. The frequency appears in Table 5.

Table 5: Results of testing the ability to correlate separated distributions

Age	Scores 1–3	4–6	7–9
8	4	16	0
10	7	8	5
12	4	9	7
14	2	8	10
TOTAL	17	41	22
		80	

The association of scores with age is significant at the 0·02 level.

Discussion

These three tests differ very considerably from each other and factors such as size and familiarity with the kind of material must play some part in determining the degree of success. However, it seems clear that, before the last year in the primary school is reached, about half the children are capable of perceiving a relationship between two generalized distributions and that this ability has been acquired by nearly all of them by the age of 14. Test F requires attention to five distributions before the best correlation can be established and is therefore a considerably harder exercise than any normally required, for example, in 'O' level map or picture interpretation. Nevertheless a quarter of the 10-year-old children and half of the 14-year-olds obtained good scores. Adults who have tried the test say that they attempt to create shapes from the dot distributions and compare them to the outlines. If so, success depends on the ability to form images of shapes, from discrete data plus a good short term visual memory and it is therefore not surprising that an association significant at the ·05 level was found between the results of Test F and those of Test B which have already been discussed.

A great deal of investigation is required into the mechanisms involved in manipulating and thinking about geographical material. In particular, the relationship between spatially based processes and verbally based processes has yet to be explored. Is spatial correlation related to a particular level of logical thought? At which stage do verbal directions assist children to make appropriate generalizations? How far can the ability to generalize at all be improved through learning? Geography appears to offer more scope for investigation along these lines than most other single subjects.

The British Journal of Educational Psychology

(Editor J. D. NISBET, Department of Education, University of Aberdeen, Scotland)

Vol. 43 Part 3 November 1973

The *British Journal of Educational Psychology* is concerned mainly with reporting the results of experimentally based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry.

The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best-known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words).

There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists. Subscriptions (£4 a year, or £1.50 per issue) should be sent to Scottish Academic Press Ltd., 25 Perth Street, Edinburgh EH3 5DW.

W. S. ANTHONY: Extraversion and Ability.

S. N. BENNETT and M. B. YOUNGMAN: Personality and Behaviour in School.

J. GARNER and M. BING: Inequalities of Teacher-Pupil Contacts.

W. YULE: Differential Prognosis of Reading Backwardness.

E. M. ANDERSON: Children with Spina Bifida and Hydrocephalus.

Other papers, and Research Notes on the Peabody Language Development Kit, the effect of reading tuition, a comparison of dyslexics and normals, and academic performance in a developing country.

BOOK REVIEWS

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Book Reviews

Integrated Studies in the Secondary School

edited by D. Warwick

University of London Press, 1973, 122 pp., £1·70
Hardback, 85p Unibook

This book is essentially an account of the ways in which several schools have approached 'Integrated Studies', together with an introduction which emphasizes the quite individual nature of each of these approaches. A teacher at each school sets the scene for his particular brand of integration and describes, with varying degrees of criticism and detail, how his own scheme has been developed. Some of these descriptions would provide useful starting-points for a school about to move into the integrated field, whilst others may yield a good crop of discussion points.

The first sentence in the book reads: 'There is, of course, no such thing as "Integrated Studies" ', and this statement is confirmed by the succeeding discussion. For me, the accounts given in succeeding chapters amount to interesting descriptions of what, and how (but not always why) particularly enthusiastic teachers have developed their integrated courses. This is not to belittle the value of their activities—I have no doubt that both their schools and pupils have been enriched by their achievements—but to indicate that no underlying philosophy of subject integration has come through sharply and well-defined in their writings.

There are no doubt powerful arguments for integrated studies, particularly in the junior forms of secondary schools—examples might be increased teacher contact-time with each class, better use of resources through team-teaching, and reduced pigeon-holing of overlapping material into traditional subject divisions. This book gives the impression that practical considerations weigh more heavily than underlying principles in 'integrated studies', even though there are probably more pages concerned with the latter than with the former.

Overall this book should be read by anyone concerned with curriculum planning or with teaching 11- to 14-year-olds and at a price of 85p it is reasonable value. But it is not a source book for integrated studies—perhaps the Schools Council and Nuffield are nearer to providing that—it is an account of the experiences of a group of teachers who have in different ways tried to remove subject boundaries, and a provocative challenge to those traditionalists who say the long-established order of separate subjects and specialist teachers must remain.

D. E. Newbold

Creative Learning in Perspective

by Tudor Powell Jones

University of London Press, 1972, 115 pp., Hardback £1·50, Unibook 80p.

Traditionally, learning and creativity would have seemed uneasy bedfellows, and today creative learning, the child of their modern educational union is enthusiastically acclaimed but still suffering from a confusion of identity. Tudor Powell Jones sets out in this book to establish a working definition of creativity with direct applicability to all children in all schools. He covers no new ground in research, and makes few innovatory suggestions concerning the curriculum. Instead he surveys the current state of research and of educational practice, marshals the evidence and presents it in a coherent form to the reader.

The writer describes research findings concerning the correlation between creativity and intelligence. He then proceeds to analyze various definitions of creativity into their component factors, drawing out from this study four main interpretations of creativity; the re-arrangement and cross-linking of elements; the resolution of mental conflict; problem formulation and solution; and creative expressiveness set free by a conducive home environment. By this means the writer arrives at a definition of creativity which he suggests can be used as a ground plan for creative learning in schools.

'Creativity is a combination of flexibility, originality and sensitivity to ideas which enables the thinker to break away from usual sequences of thought into different and productive sequences, the result of which gives satisfaction to himself and possibly to others.'

A large part of the book is devoted to a consideration of the teacher's role in fostering and developing the creative potential of pupils. There is a 30 page tour of the curriculum at infant, junior and secondary level, and from language through to

Religious Education. Not surprisingly the five or six pages devoted to each subject provide only a superficial survey of the current practice of creative learning. References to curriculum research, practical suggestions and examples from children's work are included, but the treatment lacks depth.

In defining the role of the educator, Tudor Powell Jones works from the premise that 'the school cannot create creativity, but it can do much to encourage and develop it'. He describes characteristics of highly creative people of the past, indicating that these could be a useful guide in the fostering of creativity. In fact it would be difficult to distinguish cause and results in the relationship between creativity and these personal characteristics. The unsolved problem of the creative chicken and the creative egg recurs in the chapter on measuring and predicting creativity. The writer observes that creative children, identifiable by tests of creativity, will prefer to learn creatively, though this would seem to contribute little to the general discussion. Of more concern to the teacher would be a consideration of the long-term effects of a creative education on children of low creativity, and the degree to which creative thinking can be taught. The final chapter speculates on the benefit for society of creative education, and concludes that the creative element in man is the key to personal and social harmony.

This is a short book (120 pages), and although many questions are left unanswered, the writer has succeeded in clarifying the issues so that questions may be asked more pertinently. The many pages of references and suggestions for further reading make it an excellent starting point for more detailed study.

JOY JAMES

The Learning of History

by D. G. Watts

Students Library of Education, Routledge and Kegan Paul, 1972, 117 pp., £1·75

This series is for students of education and practising teachers. Both will find this attractive little book useful. Four features immediately commend it. It is short, readable, informative and —not least in these times—encouraging.

'History has long been a criticized and controversial academic subject, but it has never been more popular.' From that paradox, Mr. Watts, as he says, sets out 'to try to resolve the dilemma it creates for teachers of history in schools'. His introductory chapter includes a summary of the sources of pressure on and criticism levelled at history teaching. In the following chapters two of these attacks are carefully examined—the writing on child psychology and development which has suggested school children's inability to understand history, and the narrow definition of it drawn by professional historians tending to exclude much of school history. In answering these two charges the author offers an account of the psychology of learning history, a definition of what we mean by history and of its characteristics, and an assessment of what it contributes to a child's development. All this will be a tonic to any teachers whose morale has been affected by recent adverse comment.

The remaining chapter translates the foregoing arguments into suggestions for syllabuses and methods. Mr. Watts appears to accept the findings of surveys of the educational preferences of adolescents that after puberty '. . . a majority lose interest and may in fact become positively hostile'. Some readers might wish he had had space to direct at these findings the critical method used earlier in the book against other propositions hostile to school history. They might also be startled by the suggestion that to meet this assumed adolescent rejection of the subject there should be a break in history teaching at 13-16, '. . . a break, filled perhaps by geography, by introducing sociology and psychology or work similar to the Schools Council/Nuffield Project'.

On the other hand as a positive approach to this problem the section on stereotypes related to a spiral history curriculum will interest teachers not familiar with this approach. The author shows that the subject can be presented differently to those pupils reacting against the history they have learned in the junior and middle school as a childish thing so it can appear more grown up. This is not necessarily achieved by introducing fresh topics which may seem more relevant to them but not more adult unless the attitude and method of study are changed. This concluding chapter also contains suggested syllabuses for junior and middle school and for the 'nadir year' of the secondary school.

There is a reading list for each chapter and a separate bibliography.

With its clear format and handy size this authoritative and up-to-date guide ought to appeal not only to students of education but also to established teachers looking for a quick means of keeping up with accepted thinking about their subject.

J. WHARMBY

Class, Codes and Control. Volume 2. Applied Studies towards a Sociology of Language

edited by B. Bernstein

Routledge and Kegan Paul, 1973, 377 pp., £4·95

Volume 1 of *Class, Codes and Control* appeared in 1971, and its successor consists largely of an empirical back-up for the theory of the earlier book. There are papers, some previously published, dealing with social class differences in the attitudes of mothers to various aspects of communication, in the characteristics of speech of children of reception-class age, and in further features of children's speech at the top infants' level. Throughout this section the style of reporting is condensed and technical so that, whatever the reward may be, staff common-room reading is not likely to be easy. A more serious consequence of the uncompromising economy is that the reader is left guessing on a number of important points of detail—for example, p. 83, how is it possible to calculate a coefficient of correlation between a geographical area and social class?

However, if the research methods are acceptable there is here abundant empirical substantiation of Bernstein's thesis. Amongst other findings it emerges that middle-class mothers place a greater emphasis on the use of language in the transmission of principles involved in skills, that the definitions they favour are explicit and abstract, and that they are less likely to avoid or evade their children's questions. The children themselves show social class distinctions in the degree of differentiation of their speech, in their expression of uncertainty, in their definition of control situations, and in their fluency of utterance (though sex and hesitation position effects complicate the issue here). In each case the social class difference is in the predicted direction.

Are the methods acceptable? Critics of Bernstein are unlikely to be persuaded that all has been confirmed. The familiar objections are open for use that maternal reports are not the same thing as maternal behaviour and that children's speech reactions to pictures and story-telling requests may be strongly influenced by the material and social conditions of the situation. But this is a difficult area for research, and the authors are generally aware of the limitations in technique. What a pity they, and the editor, do not seem to have been equally conscious of the off-putting effect of stern and unrelieved technicality!

Fortunately, intelligibility is established in the final part of the book where two theoretical issues are extensively discussed. Ruqaiya Hasan writes on 'Code, register and social dialect', and J. A. Cook on 'Language and socialisation'. Both essays are immensely useful in their integrated survey of basic concepts, and can be thoroughly recommended to the relatively unsophisticated reader. The book concludes with an Appendix on 'The functional basis of language' by Professor M. A. H. Halliday, who also contributes the, to me somewhat suspect, 'Introduction'.

D. J. BRUCE

Reading for Meaning

by Pat D'Arcy

Hutchinson Educational for the Schools Council, 1973

Volume I: *Learning to Read*, 307 pp., £3·50
Volume II: *The Reader's Response*, 265 pp., £3·50

Reading For Meaning is a two-volume review of current practices and research in reading. It is a work of selection and summarization of central topics, and should be judged as such. The style is non-technical and the chapters are meant as bases for discussion. Some readers may find the sheer weight of familiar material, particularly in Volume I makes straight-through reading somewhat dry.

This volume, *Learning to Read*, deals with major issues in the initial teaching of reading and would be useful as a reference source or text-book for those requiring a digest of information. In addition to such predictable topics as 'reading readiness', reading schemes, teaching methods and devices, and the influence of linguistic factors, there is an interesting chapter on the pre-reading period in which attitudes towards books and the capacity to respond to literature begin to form. This is a theme which is taken up again more fully in Volume II.

The first volume is disappointing however in that Mrs. D'Arcy seems to become overwhelmed by the factual density of her material and half-way through the book the style of summary with some evaluation gives way to sheer précis. This second part is a series of appendices in which major researches, descriptions of reading schemes, summaries of important theories and a list of suitable infant readers are all presented in catalogued or edited form.

This bland and straightforward mode of presentation might be a virtue if the book is to be used as a college course book—a use to which it would be well-suited but for its price—or by busy heads of schools and departments who wish to become generally informed or fill gaps in their knowledge. However, given that much of the factual information is in itself unexceptional and could easily be obtained from other sources—the description of Stott's Programmed Reading Kit reads very much like the publisher's information leaflet—so one would welcome a more intrusive authorship. As an English specialist, Mrs. D'Arcy could have added a fresh perspective to a field that has usually been described from the point of view of the psychologist or infant reading teacher.

One wonders how accurate much of the material is in detail; the list of 'reading readiness' tests is rather out of date and misleading. In fact this is the only point at which reading tests are dealt with. There is no discussion of the materials and techniques currently employed in assessment of reading attainment or diagnosing reading difficuty. This is a subject which causes many teachers concern and vexation and its omission is disappointing. Reading tests are overdue for scrutiny from the point of view of English teachers and Mrs. D'Arcy missed an enviable opportunity in this respect.

In the second volume, *The Reader's Response*, Mrs. D'Arcy discusses what might be termed the secondary aspects of reading. The notion that reading is a continuous process, and that there are reading skills to be acquired at secondary school which are just as much 'reading' as the initial skills, has not been appreciated in the past. This volume, however, provides the secondary and older junior teacher with a valuable introduction to the cognitive and affective aspects of reading once the initial decoding process has been mastered

A number of studies of children's reading habits and attitudes are summarized. Certain clear determinants of preference such as sex-differences, academic qualities, and age appear consistently in the research, but there is little in the findings to surprise any secondary teacher of English. The gulf between the literature favoured by children in their voluntary reading and that approved by teachers as worthwhile is clearly indicated. It is most characteristically stated in Connie Alderson's 1968 study of teenage magazines. References are made to the perhaps dismaying popularity of Enid Blyton and Ian Fleming and there are certainly a number of researches, not reported in the book, such as that carried out by ACE in 1968, which support these findings. At the same time one wonders whether there has not been a recent hardening in children's reading tastes: Flemming has been displaced by 'Richard Allen', author of *Skinhead* and its sequels, and seems almost genteel and literary by comparison.

An approach given at least cautious approval in the book is that 'inferior' reading should be tolerated because readers tend to become more discriminating and sophisticated with age. Furthermore, reluctant readers may find the enforcement of reading standards in their choice of reading has an aversive effect. The argument is that even the most worthless reading elicits a response which is similar in kind and continuous with that elicited by worthwhile reading, only the quality is different. Many teachers, faced for the first time with the negative or evasive attitudes of lower stream pupils will wish for a much fuller account of research and practices in this field. In fact the amount of information about *changing* tastes is limited, even though the need for such work is clearly indicated by the pattern of taste revealed in surveys. A. K. Pugh has observed that children do read much more than teachers assume, even though the quality of material may be poor. The book reviews evidence which supports this, but it does not continue to a further point made by Pugh; that teachers over-estimate the level at which pupils are capable of reading. Nor does it include any mention of studies like that of Kenyon Calthrop in which successful use of 'approved' literature in the classroom is investigated. Thus teachers who are facing on a day-to-day basis the uncomfortable fact that there are many children who whole-heartedly reject books and reading, or, whatever their capabilities, only seem to enjoy *Skinhead* and *Chopper*, will get but limited guidance from this part.

There is much more positive information in the section on reading response in the form of comprehension and appreciation of literature. The arguments between a multi-skill and unitary theory of comprehension are discussed and the author is inclined to favour the latter as a basis for assessment. A number of little-known but intriguing studies of critical response in secondary and university students are reported. These studies, carried out by such diverse researchers as Eysenck and James Britton, will be of special interest to the secondary teacher of English. They deal with the specific difficulties and sources of confusion together with idiosyncracies of linking which shape or distort the readers response to literature. This research resulted in the identification and description of many sources of aberration which teachers could well consider in their own pupils. They have also revealed that the student's and pupil's capacity to respond, both at secondary and university level, often falls short of the teacher's assumed or expected level: grammar school children may base their preference on novels or relatively crude criteria, while undergraduates may be very unsure of themselves in voicing preference for a particular poem, or may have difficulty in discriminating between genuine literary creations and pseudo-poems devised for experimental purposes. However, the researches also draw a related conclusion, which is that given *time* the capacity to respond confidently or discriminate becomes surer and that taste and attitude become more discerning.

This second volume successfully draws attention to the secondary teacher's responsibility in the teaching of reading. It shows there are both tasks and problems to be faced which are just as real as those associated with reading in the infant class.

There is also a better balance between description and comment in this volume and one senses something of the author's own outlook, particularly in the way she gives prominence to work done by groups and individuals associated with the National Association for the Teaching of English.

Reading for Meaning may well play a useful role for teachers who are becoming more aware of their responsibility for literacy and feel the need to assimilate knowledge and information about the field. More specialized readers will be disappointed by the absence of new or unfamiliar material, particularly in the first volume, while the more hard-pressed secondary teacher may feel the second volume fails to dramatize his problems as a teacher of literature. This second volume also leads one to conclude that the area has been sadly under-researched, although one hopes work in progress by NATE and the Schools Council will do something to rectify this.

DENIS VINCENT

Social Control and Socialization: a study of class differences in the language of maternal control

by Jenny Cook-Gumperz

Routledge and Kegan Paul, 1973, 290 pp., £4·25

Dr. Cook-Gumperz's book is the sixth in a monograph series edited by Professor Basil Bernstein, emanating from his Sociological Research Unit (SRU) at the Insitute of Education, and dealing with various aspects of his thesis that the poor school performance of working-class (WC) children can be explained in terms of a class-based language deficit. Her book, which is an abridged version of a doctoral dissertation carried out under Bernstein's direction, reports an extensive investigation of a central tenet of his thesis, namely, that a child's home linguistic environment is responsible (or, to blame) for these deficits; and, in particular, his opinion that the strategies of maternal control are reflected in the sort of language a mother uses, and these strategies will differ according to class.

Dr. Cook-Gumperz administered a questionnaire to mothers of a year's intake at 15 primary schools with WC catchment areas and 5 schools with middle-class (MC) catchment areas; the questionnaire was concerned with how they controlled their child. A similar questionnaire was administered to the 5-year-old children as to how a mother might control her child in a variety of hypothetical situations. An elaborate grid was devised with help of colleagues at the SRU to code the replies according to Bernstein's categories of social control—imperative, personal and positional—and the extent to which control depended on explanation (on appeals to reason, appeals to emotion) or on punishment. As expected, there were class differences: WC mothers said that they tended to use more imperative techniques which depended little on explanations; MC mothers said they were more likely to use strategies directed at personal aspects of the situation, and they would make more appeals to reason. A criticism levelled directly at Bernstein by Rosen, and indirectly by Labov, is that he does not account for interview artefacts. It has been shown in other situations that interviewees respond differently according to their class and the perceived class of the interviewers, and it is entirely possible that the data accumulated here reflects not the real class differences in control strategies, but class differences in interview responses by W and MC mothers to the MC social science students who were conducting the interviews. Perhaps, MC mothers were more conscious of what they thought were socially desirable control strategies than WC mothers, and answered accordingly. This is borne out by her finding that the children (who would have such preconceptions, yet who presumably suffered different controls) did not show class differences in this respect.

An even more thorny problem concerns the application of Bernstein's celebrated concepts of 'restricted' and 'elaborated' language 'codes'. In an account of these notions which is, understandably, entirely uncritical, and, less defensibly, makes no suggestion that anyone has ever criticized them, Dr. Cook-Gumperz offers at least two mutually contradictory theoretical definitions ('codes' are sets of linguistic preference rules presumably acting on pre-existing language abilities, and at the same time, they are these language abilities) and a third, operational, criterion derived from the textual component of Halliday's and Hasan's systematic grammar. This is not the place for a theoretical discussion of language codes, but two points about their application here are worth making. First, the correlation she finds between control strategies and language may be explained not in terms of intrinsic class/code differences, but in terms of the well-understood fact that people will talk about different things in different ways; and naturally, one can detect syntactic, as well as vocabulary, differences in 'Don't do it' and 'if you paint our new Habitat sofa, Justin dear, it will clash with the cushions'. Second, the study she made of the children's speech was, unfortunately, the part of her dissertation cut out of the book.

Dillon's University Bookshop

Our Education and Psychology shop
carries a large and comprehensive range
of titles on these subjects.

Orders can be dealt with by post, but why
not call in and consult our friendly and
knowledgeable staff?

Our catalogue of new books in Education
and Psychology is now available: let us
have your name and address and we will
send you a copy.

Dillon's University Bookshop Ltd
1 Malet Street London WC1E 7JB

Her report reads like a doctoral thesis: it is very detailed and thorough in the presentation of the techniques and results, and a wide range of speculations are included in the theoretical sections, but in fact, she suffers from Bernstein's fault of going much farther than the evidence permits.

BRIAN BUTTERWORTH

A Language Program for the Non-language Child

by Burt Gray and Bruce Ryan

Research Press, Champaign, Illinois, 1973, 181 pp., $5·95

The term non-language child implies purely 'a non-performer of the verbal-linguistic code'—the label, autistic, receptive aphasic, dysphasic, language delayed, brain damaged, etc. 'makes no difference as the teaching job is still the same'. Thus this book has relevance to speech therapists, teachers of specifically handicapped children and mentally handicapped alike.

The problem is 'the child doesn't talk much' and traditional education errs in assuming language so that these children suffer. The training approach advocated by the authors is behaviour modification. The basic learning sequence is *conditioning*. The book thus comes at the right time for only this year Professor Mittler, Head of the Hester Adrian Research Centre at Manchester University, who also believes in behaviour modification techniques, argued that a group of educational psychologists should go to the United States to learn these techniques and come back as the apex of a 'pyramid' structure of teaching in the field. 'Pyramid Selling' is, of course, another transatlantic import.

Well, this book sets out a language programme which may save such a journey or pyramid group. Any reader is given explicit instructions in the stimulus-response-consequation sequence of *programmed conditioning*. Phonology, morphology, and grammar are each handled. The child may begin with very simple responses working up to say, 'the ball is red'. Higher levels are reached by increasing the length and complexity of responses through 240 steps. Reinforcers (commonly tokens, styrofoam stars, chips, beans, etc) are given to the child for responses. These are exchanged for toys. Another form of reinforcer is the utterance 'good' which is said after every correct response.

The sample programme illustrating the approach indicates a beginning in the identification of nouns, naming nouns, using *is/on*, *'what is'*, *he/she/it*, *I am*, singular nouns in the present tense, plural nouns in the present tense and so on.

Work sheets are recommended for the essential minute recording of progress. In addition a Programmed Conditioning for Language Test (PCLT) has been devised indicating the progress of the child through the programmes, and an adequacy score to show how effective expressive language has become. Criterion reference tests are also recommended to assess the progress of the child before and after programmes. Spontaneous language should also be recorded. Home Carry Over Record forms indicate the work to be done at home under the supervision of parents, five to ten minutes a day is recommended.

A reasonable portion of the book is devoted to the authors' experience with the methods with pre-school children using programmed conditioning sessions of 90 minutes. The remainder of the time was used for less formal situations. Sessions were also conducted with mentally retarded and hearing impaired children. The experience of the authors is regarded by them as 'quite promising'.

We are all searching for ways to help children who cannot talk, or read or write, whatever their label and whatever the cause. There is no doubt a place for behaviour modification techniques. Perhaps I lay myself open to being regarded as old-fashioned if I find manipulative ideology repugnant. For those who wish to improve their professional expertise in the field of conditioning techniques this is a very clear and expertly written book.

J. E. Collins

A Place for Learning: aspects of secondary education in three English-speaking countries

by Frances Stevens

Hutchinson Educational, 1972, 316 pp., £6·50

This is a comparative study of some aspects of secondary education in three countries: England, USA and Australia. Frances Stevens sets out to explore in the last two countries the relationship between social background and school success, pupils' experience of schooling, their life outside the classroom, their relationships with each other, the reasons for under-achievement, how they were viewed by their teachers and, correspondingly, how they viewed those who taught them. An observation sample of eleven schools in Illinois (excluding Chicago) and two schools in San Francisco, and a variety of secondary schools in South Australia was chosen. The method of investigation was by means of structured interviews, discussion with pupils, teachers and educationists, scrutiny of school records, questionnaires and sociograms. In the main the pupils studied were aged between thirteen and fifteen. The data obtained abroad were then systematically com-

FAMILY HISTORY IN SCHOOLS

by **Donald Steel** and **Lawrence Taylor**

The first volume in a new series on the teaching of environmental history, based on the authors' original and nationally-known experimental work at Berkshire College of Education.

It propounds, with practical guidance on method, an entirely new concept in history teaching, whereby the child starts with himself, and conducts his own research into his immediate family.

200 pages, 83 plates, limp covers £1.50

PHILLIMORE
SHOPWYKE HALL
CHICHESTER

pared with similar information previously collected by the author on secondary pupils in England. The technique followed was largely impressionistic and the research design did not require statistical procedures except at a very elementary level. The author, indeed, lays no claim to absolute objectivity—in fact, the interpretative nature of her findings is one of the most attractive features of the study.

These findings confirm a certain sociological universality, at least as regards industrialized societies. Educational achievement in all three countries is demonstrated to relate to variables such as social class, race, religion, neighbourhood, friendships within peer groups, family size and the emotional 'climate' of the home. As regards the curriculum, there is general agreement that English and mathematics are the most important subjects—although the author speculates why this need be so. Interestingly, the ideal teacher is seen in almost identical terms by children everywhere: the adult who can communicate and explain, exercises patience and care, establishes a participatory dialogue with his class, presents his lessons in a variety of ways and is relevant, and has an agreeable delivery. The 'golden years' of school in all three countries would seem to be about the ages of eleven to thirteen—presumably before the pressure of examinations and the preoccupations of a future career have begun to build up.

Some shrewd incidental comments interlard these conclusions: the anxiety, almost to the point of neurosis, to which teachers are prone—even the most competent; a rigorous criticism of the behaviourist approach to learning, which may tend to dehumanize the individual. As the author of *The Living Tradition* and *The New Inheritors*, with a wide and varied experience behind her, in this new study Frances Stevens merits careful consideration when she speaks of the need for research into 'the relationship of feeling to achievement, the association of affection with learning, the relationship between pleasure and learning, and the connexions between friendship and learning'. The affective aspects in the process of learning have been indeed too much neglected in the past.

The book is written in that easy interesting style we have come to expect from a cultured scholar. The results are summarized in simple tabulations. The case histories render the presentation vivid and easy to follow. For one reader, however, the nub of the work lay in the challenging statement halfway through: 'Few . . . have considered the curricular and organizational implications of putting social unity in the forefront of our educational aims. Does anyone seriously wish to do so? If so, he must initiate a revolution.' After the airy theorizing and pseudo-findings of so much that passes for educational research, this comparative study, unambitious in its design yet penetrating in its results, comes as a refreshing breeze of common sense.

W. D. HALLS

Toward an Educational Theory for the Mentally Handicapped

by John McG. McMaster

E. Arnold, 1973, 134 pp., £2·00

In this book McMaster attempts to lay the foundations of an educational theory for the mentally handicapped, incorporating ideas and concepts from a number of key areas. The importance of this task springs from his belief that the 'greatest requirement of the present time is a conceptual framework which will allow a holistic view of a fragmented, unstructured and largely chaotic field'. Describing educational theory as 'practical theory', primarily because of the inclusion of 'shop floor' experience as a vital element, he attempts to demonstrate how history, psychology, sociology, medicine, psychiatry, political, economic and social studies each contribute towards a theory from which aims and methods can be deduced. Such an approach is welcome for at least two reasons. Firstly, as responsibility for the education of the severely subnormal has recently been transferred to the DES, the time is opportune for making a fresh start in developing a more realistic, systematic and coordinated approach toward meeting the educational needs of these children. Secondly, a holistic view is essential if we are to achieve a complete understanding of the educational process, rather than rest content with the partial explanations of any one discipline.

In view of this, it is a pity that, after reading the book, one remains unconvinced that it contains anything other than a hotchpotch of material from a variety of sources, revealing no consistent attempt to evolve a unified conceptual structure or reorganize the field in a way which might elicit new insights. Unfortunately, the seeds for this appear to be sown in the second chapter where educational theory is described as essentially eclectic in nature: once this is assumed, any attempt to unify and integrate is by definition impossible.

Take for example the contributions from psychology and sociology. Clearly there is an overlap of subject matter here, but the point is how are they to be fused into a practical theory? In the chapter on sociology we read 'these problems (i.e. of child and family in a social context) are extremely complex and have a *direct* effect upon the child, his development and his education' (my italics). This seems reasonable enough. But just how important is it to take account of these factors, when, from psychological considerations, we decide to make use of a behaviour modification programme based on operant conditioning techniques? The four-and-a-half-year-old with a social behaviour problem may have responded well to this kind of psychological approach, but to what extent would knowledge of sociological factors have influenced our decision to use such techniques? In such situations, McMaster provides no rationale to guide us, and it is precisely because we need a rationale that it is relevant to develop an educational theory.

Central to the task of formulating an educational theory must be an attempt to make as clear as possible the basic differences between schools of thought within the various subject disciplines. It is particularly important when the disciplines concerned include human studies for a writer continually to 'bare his soul' on certain fundamental issues. Although McMaster does, in passing, refer to controversies he fails to mention some of the important implications of these. For example, one of the main methodological problems in sociology is not that sociology appropriates concepts the layman uses in everyday speech, but arises from the existence of basic disagreements among layment and sociologists alike about the nature of man and the world. Opponents of behavioural science point out that many so-called objective studies are based on extremely subjective assumptions; and in sociology and psychology the problem of objectivity has yet to be solved. (It is this reviewer's opinion that many studies of families with mentally handicapped children have been unproductive and inconsequential mainly because they were in the worst traditions of empiricist sociology.)

In his discussion of the historical contribution McMaster misses another opportunity to mention important alternative approaches. Again, he starts well with a telling quote from Simon (1966): 'There is perhaps no more liberating influence than the knowledge that things have not always been as they are and need not remain so . . .' Then we travel a familiar road—the contributions of Montessori, Galton, Itard and the Wild Boy, Seguin, Binet, the various Education Acts and committee reports; in fact, a history of certain psychological theories, educational practices and state involvement in the education of the mentally handicapped, with very little reference to the 'societal backcloth' mentioned by McMaster himself on page 41. We learn nothing about the day-to-day existence of a family with a mentally handicapped child; what problems they face and how they cope with them; or of the attitudes of ordinary people towards the mentally handicapped. With reference to the last point, McMaster's position is clear:

> It may appear, in retrospect, that the strong emphasis in the nineteenth century upon placing mentally defective persons in residential institutions, rather than attempting day provisions of an educational nature, was the result of a general attitude which was hostile to any form of integration with society. Probably the populace at the time would have adopted such an attitude, still prevalent today; however, it is clear that from administrative, legal, medical and philanthropic viewpoints the segregation advocated was thought to be in the interests of the mentally defective.

There are two points to note here: the implicit acceptance of the idea of gradual enlightenment spearheaded by middle-class professionals; and the assumption that the populace has always held unenlightened attitudes. Yet it might well be argued that the mentally handicapped were segregated more as a result of a reaction by the forces of 'law and order' than because of pressure from the populace (witness the 'putting away' of unmarried mothers and other offenders against establishment morality) and that where family life had not broken down the handicapped child often functioned as an accepted member of the community where he lived, e.g. the 'village idiot'—an idiot, yes, but still part of the village. It is also worth noting that rather than pressing for segregation, working-class (as well as middle-class) parents often resist pressure from the authorities to transfer their handicapped child to a special school.

In conclusion, one should consider the relevance of this book to teachers. There is no doubt that McMaster always has the teacher in mind as he incorporates feedback from the teacher into his theory. He recognizes the importance of teachers stating aims systematically and analyzing their methods. In this respect, his work is typical of the current concern with the 'systematic approach to educational objectives'. This is all well and good, but the question nags—which educational theory and who is to decide upon aims and methods in any one school? Schools, particularly special schools employing behaviour modification techniques, do not easily lend themselves to a variety of different regimes under the one roof. It will obviously be important for *all* members of staff to be involved in planning at all levels, i.e. at the 'theory' as well as 'aims and methods' end, and to have a real say in the implementation of policy. It is likely that a bad theory could prove more restricting and damaging than no theory at all, and teachers should make sure that procedures exist whereby a bad policy can be swiftly reversed, if this is felt to be necessary.

J. C. Quicke

Sound and Sign: Childhood Deafness and Mental Health

edited by Hilde S. Schlesinger and
Kathryn P. Meadow
University of California Press, 265 pp., £4·50

The paucity of relevant works of research concerning the education of deaf children tends to make the co-worker in this demanding field of endeavour more ready to express gratitude than to adopt the critical attitude appropriate to a reviewer—especially where the research in question is as capable and thorough as that of Professor Schlesinger and Dr. Meadow. Another consequence is that single works of research come to be regarded as conclusive rather than corroborative or otherwise of other studies. Professor Schlesinger however claims properly that these studies should provide the basis and impetus for additional research, as one can have little doubt they will.

What is most impressive about this book is the emphasis throughout on communication not as something with which the parent and teacher aim at endowing the deaf child, but as an integral part of the biological development of the child; not as an end in itself but as a means whereby the child adjusts to his social and intellectual environment and whereby emotional development proceeds. The thesis that these purposes are facilitated where oral communication is supported and expanded by manual communication is persuasively argued.

One's main reservation in welcoming this book is a regret that no mention or discussion is made of the basic difference of principle seen by Tervoort as distinguishing the communication system built of acoustic units (oral) from one built of visual units (manual) nor of whether such disparate systems may be truly compatible. Intensive study of the various manual systems of communication is required, perhaps by linguistic analysis (although deaf education is as short of linguistics as of other specialists) in order to determine which type of manual system is most suited not merely to coexist with oral development, but actually to nourish and foster it.

A second reservation, by no means confined to the book under review, is that the cases of individual deaf children chosen for the purpose of illustration are atypical in the matter of the favourable factors operative, and thus an over-optimistic climate of thought engendered, albeit not deliberately, which may cloud the reader's judgement. Most teachers of the deaf in the reviewer's own area would read with considerable envy of the degree of insightful involvement shown in the parent-child interactions quoted.

These reservations apart, the book must be required reading for all who work with deaf children. In particular, the beautiful quotation from Montaigne that 'The word is half his that speaks and half his that hears it' is one which all exponents of 'auditory training' would do well to ponder.

G. M. Fraser

The Free School

by W. Kenneth Richmond
Methuen, 1973, 211 pp., Hardback £2·90,
Paperback £1·40

In *The Free School*, W. Kenneth Richmond takes a quick canter through the literature of the de-schoolers (Roszak, Reimer, Illich etc.), relates their arguments to the British situation, accepts them more or less wholeheartedly, takes a look at the depressingly few examples of radical theory put into practice—and that's it. Into hard covers, published at £2·90, it's another solid brick in the wall of words that securely protects so many academics from the nasty cold world in which people actually try to change things.

If 'radical' academics and educationists continue to lift the pen rather than the hand to the cause of real educational reform, they must accept that they are in fact only defending their own status and the *status quo*—the two after all go hand in hand.

As the author says on the first page, the underlying assumption in most of the fashionable discourse on de-schooling and free schooling is that it is not to be taken seriously: 'To affect a bogus radicalism suits the mood of the moment'. There must have been quite a race to produce the first bumper amalgam of the literature, to bag the obvious title.

Bitter? Well, yes. Richmond has obviously carefully and critically read (and laboriously footnoted) most of the current critical literature that makes up the rag-bag title of the de-schooling movement. He does not quibble with, in fact he accepts and in some cases restates with more force most of the basic arguments (schooling is not synonomous with education, compulsion is incompatible with education etc.). But the result of all this commendable intellectual effort is just another book.

Oh well, this is just another review, so I might as well recommend the book as an introduction to the real literature, in the hope that it won't be taken as an excuse by all those bogus radical students for not reading the original works—or even worse—as some sort of absolution from radical action.

Take three paragraphs.

> As applied in the past, legal compulsion may have served a useful purpose. In a less servile society, its retention is as anachoronistic as that of corporal punishment. Until the education system is rid of it, schooling cannot be other than illiberal.

And:

> From start to finish, schooling is essentially élitist, grading and classifying pupils, favouring those who do well in its eyes, demoting those who do not. Though everyone is on the ladder (always supposing that it is one and the same ladder), it is obvious that its higher rungs are only attainable by climbing over the backs of those who are at the bottom of the pile . . . the school, as a selective agency, makes nonsense of the concept of equality of opportunity.

And:

> The school's attempt to put all its eggs in one basket is misconceived. The days of its enclosure are numbered. Very soon there will be no point in its trying, and less in pretending to believe that learning can be 'cabin'd, cribb'd, confin'd' in this way.

If these are accepted by the reader as they are by the author, they imply absolutely basic changes in our conception of the education system. That means rather more than a bit of curriculum reform here, an extended day and open-plan building there.

The author has scraped the barrel in his attempt to find examples of 'de-schooling' in practice—the Open University might reasonably be surprised to find itself featured along with a couple of community-minded comprehensive schools. If Mr. Richmond had stayed his pen a little longer he might have found some more realistic examples of free-schooling at least: de-schooling is, after all, still not a legal option.

PETER NEWELL

SRHE Publication

One of the more recent SRHE publications is **Register of Research into Higher Education in Western Europe 1973**, which lists on-going researches in this field under a number of subject areas. It gives details of the aims of each research, those engaged on it, the methods used, any interim results, and references to any publications so far produced. The Register is available from the Society for Research into Higher Education, price £3·00.

The NFER Publishing Company Limited – Test Division

The principal activities of the Test Division of the NFER Publishing Company are the publication and distribution of British and overseas psychological tests and associated materials.

New Publications from the Technical Department's Research and Standardisation Programme

British Manual to the Bennett Mechanical Comprehension Test, Forms S and T *by Peter Saville*
Normative, reliability and validity data for a large number of craft and technician apprentices in commercial and industrial organisations and colleges. Test Range: 15 years +.

The Gates MacGinitie Reading Tests, Primary A Form 1 British Supplement of Norms *by Peter Saville and Stephen Blinkhorn*
British norms are based on over four and a half thousand primary school children between 6·10 and 8·00 years. Facility values and reliability of the tests are also considered. Test Range: 6·00–8·00.

British Manual to the Mechanical Comprehension Test Form CC *by Peter Saville*
Form CC of the Mechanical Comprehension Tests is for higher engineering ability levels and this manual contains normative, reliability and validity data for ONC, degree and military groups. Test Range: 16 years +.

British Supplement to the High School Personality Questionnaire *by Peter Saville and Laura Finlayson*
Norms are presented for a representative sample of boys and girls (13–15 years) at schools in England, Scotland and Wales. The equivalence of the first and second order factors for Forms A and B, cross-cultural comparisons between British and American children, and inter-factor correlations are also considered. Test Range: 13–18 years.

New Tests and Other Publications from the Test Division

AH4 Machine Scorable Edition *by A. W. Heim (Test) and Peter Saville and Janice Hare (Users' Guide)*
A new machine scorable version of Dr Heim's well-known test of general ability has been devised in conjunction with the Technical Department. Test Range: 11 years +.

McCarthy Scales of Children's Abilities *by Dorothea McCarthy*
A new individual test of mental ability yielding six scores: Verbal, Perceptual Performance, Quantitative, General Cognitive, Memory and Motor. Test Range: 2½–8½ years.

Object Relations Technique—Introduction *by H. Phillipson*
A new introduction to this projective technique now replaces the original 1955 text. Test Range: 11 years +.

Panda Test and Stycar Vision Test Manual *by Mary D. Sheridan*
The Panda Test is a recent addition to the Stycar Vision Tests for severely visually handicapped young children. The Stycar Vision Test manual has been revised to include information on the next test and its development.

Test of Motor Impairment *by D. H. Stott, F. A. Moyes and S. E. Henderson*
A new clinical instrument originally intended as a review of the Oseretsky Test but which, after six years of development, has emerged as a virtually new technique. Test Range: Sub 5–13 + years.

Further details and information:
NFER Publishing Company, Test Division,
2 Jennings Buildings, Thames Avenue, Windsor, Berks., SL4 1QS

chambers

Young Set Dictionaries

Amy L. Brown, John Downing and John Sceats

for each primary school child–the right dictionary

A unique set of graded dictionaries. Illustrations, which constitute *Dictionary One*, are gradually phased out in *Two*, *Three* and *Four*. In these, however, the definitions are expressed in the language of children—simple, colloquial and to the point.

'conceivably the most useful and usable children's dictionaries ever produced.' *The Times Educational Supplement*

Dictionary One 40p
Dictionary Two 40p
Dictionary Three 50p
Dictionary Four 60p

Inspection copies from
Chambers 11 Thistle Street Edinburgh EH2 1DG

EDUCATION 3–13

A New Journal of Primary Education

Published by Collins Educational and the Primary Schools Research and Development Group, University of Birmingham

Published April and October

Annual Subscription £1, Post Free

Available from:

Education 3-13, Collins Educational,
Kirkintilloch Road, Bishopbriggs,
Glasgow G64 2PW

RESEARCH IN EDUCATION

November 1973 *Number 10*

A limited empirical study of the attitudes of serving teachers to educational theory: PHILIP H. TAYLOR

Coalition formation in colleges of education and the introduction of the B.Ed.: J. NIAS

Adjustment in the classroom 1. Teachers' ratings of deviant behaviour: J. A. WILSON

Role orientations and sub-cultures among undergraduate students: LOUIS COHEN and DEREK TOOMEY

The effects of pre-test difficulty on post-test performance following self-instruction: JAMES HARTLEY

University course-changers and potential course-changers: MICHAEL P. HORNSBY-SMITH, CELIA J. NEWBERY and SALLY R. HART

Reviews

by P. E. Vernon, James Drever, Neil Warren, Philip Williams, William Lamont, Louis Cohen, K. Lovell and Ernest Rudd.

Published twice-yearly, May and November

MANCHESTER UNIVERSITY PRESS

NFER COMPUTER SERVICE

The National Foundation for Educational Research has now installed its own computer—an IBM 1130.

We anticipate a certain amount of spare time in the next year or so, and are prepared to make the computer available to Educational Research Workers at favourable rates.

Card-punching will also be undertaken, at competitive rates.

Details

A comprehensive range of programs is available. These include regression analysis, factor analysis (up to 100 variables), a general tabulation program, χ^2 analysis, and analysis of variance. In addition, there is a wide range of data processing programs for those researchers who need to handle a large amount of data.

Rates

Computer: £20 per hour.

Data Processing: £24 per 1,000 cards. Punched and Verified.

Specimen Costs:

Regression analysis of 8 variables on 40 subjects—£5: 15 tables from 2,600 subjects—£10.

All charges subject to 10% VAT.

Telephone Slough 28161, or write to

A. N. James,

NFER, The Mere, Upton Park, Slough, Bucks.

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £2·00 post free. Overseas (excluding USA*): £2·30 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: 75p post free. Overseas*: 85p post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Enquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Bucks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the text, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or
2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© *National Foundation for Educational Research in England and Wales*, 1973

Designed by H. A. Senter

Printed by King, Thorne & Stace Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

12/8/75

In this issue

Subject choice in secondary school

Reading schemes and vocabulary

Three myths in educational research

ARC
Docket & issue
to the librarian
R. Das
7/8/75

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd.

EDUCATIONAL RESEARCH

Volume 16 Number 2
February 1974

Contents

A developmental study of the acquisition of some moral concepts in children aged 7 to 15
J. B. Edwards 83

Religious beliefs and moral values: the influence of the school *J. Marvell* 94

Religious socialization in the junior school *P. Jarvis* 100

Drugs education: questions and answers *D. A. Lane* 107

In search of an 'early teaching grammar' *J. L. Presland* 112

Scoring written work in English as a second language: the use of the T-unit *C. Gipps and E. Ewen* 121

Creativity and curriculum structure *E. Ogilvie* 126

Linked courses in Herefordshire *D. H. O'Donnell* 133

'O' level or CSE? The consequences of double entry *J. Fry* 139

Question choice in examinations: an experiment in geography and science
E. G. Taylor and D. L. Nuttall 143

Book reviews 151

Diary No. 908.(3).
Date 7. 8. 75
Bureau Edu. Psy. Research.

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 16 Number 2
February 1974

A Developmental Study of the Acquisition of some Moral Concepts in Children Aged 7 to 15

J. B. Edwards, Educational Psychologist, North Wales

Summary

An investigation was carried out to ascertain how 700 children aged seven to 15 (half boys and half girls), acquired the ability to define certain moral concepts, how they decided on whether an action was right or wrong, and what influenced the acquisitional process. The results showed older age-groups gaining significantly higher mean scores in defining concepts, with a sex difference favouring girls, but no significant social class differences. However, even older age-groups encountered difficulties with some concepts, e.g. 'doing right' and 'being good'. In deciding whether an action was right or wrong, more older children claimed they decided for themselves, and a significantly higher number of older girls stated that they did so. The sample claimed that their mothers had been a major influence in acquiring moral concepts, much more so than peer-groups. The mass media was not thought by the sample to be accepted uncritically, and the church was believed to help them most. Evidence from this research would suggest that the teacher's contribution, to the acquisition of moral concepts during the early school years, seems highly relevant.

Introduction

John Wilson (1972) has asked whether Piaget—and many other psychologists for that matter—have failed to distinguish between a child 'having' a concept and 'being able' and 'willing' to use a concept. Wilson, as an example, believed one should take language more seriously than Piaget did, and probe deeper into what a child 'meant' when he said 'right', 'ought', etc. in research carried out on moral judgment. Moreover, one would also like to have an unambiguous picture of just how adults normally manipulated such concepts. Hence Wilson's view confirmed that of Peters (1969) when he believed that children have to learn to use concepts such as 'right', 'wrong' and 'ought'; and that moral education should aid children 'to get on the inside of the family of moral concepts which structure our social world'. Piaget (1932), when asking young children in his sample questions which included such words as 'punish', 'naughtiest', 'right', 'good', 'revenge', remarked that on some occasions the children did not appear to comprehend these terms.

In order that some further light be shed on this problem, it was felt that (a) children could be invited to describe how they defined some of the basic terms used in moral discourse; (b) how they made up their minds between 'right' and 'wrong'; and (c) what they believed had been influential in this process of acquiring moral concepts. However, some previous studies of an empirical nature would first have to be mentioned.

Some Relevant Empirical Studies

Language development

Vocabulary lists from individual intelligence tests, e.g. the Wechsler Intelligence Scale for Children can indicate the age levels at which children are able to adequately define words presented to them; and abstract words, e.g.

'revenge', 'obedience', are introduced on the revised Stanford-Binet test at year 11 level. Thus any vocabulary score is closely related to verbal ability as assessed by the previously mentioned tests, and also age is related to increased mean vocabulary scores. Furthermore, sex differences in favour of girls have been reported by a number of studies. Macarthy (1956), in a long review article on sex differences in language, was of the firm belief that where groups of boys and girls were well matched in intelligence and social class background, and when the situation in which responses were recorded did not tend to favour the interest of either sex, 'there appear slight differences in favour of girls'. Terman and Tyler (1956) also thought that most of the many studies based on school achievement tests have shown girls to be superior in tasks involving the use of language and verbal fluency. More recently in a classified summary of research into sex differences in verbal abilities (Maccoby, 1967, pp. 334-8), out of 24 reported studies on children's verbal fluency and vocabulary, girls gained superior scores in 11 studies and boys in six, with seven studies not showing significant differences.

Obviously, such findings would be relevant to any attempt to define words used in a moral context.

(b) *Evaluation of 'right' and 'wrong'*

Coincidental with the increase in ability to define words commonly used in moral discourse, is the question of whether children are able to decide more for themselves whether certain courses of action are 'right' or 'wrong' ones. Do children, for example, come to rely less on external authorities and depend more on 'self-evaluation' as they progress through adolescence? Aronfreed (1969), in a comprehensive review of empirical studies dealing with what he described as 'internalization', stated that one of the remarkable consequences of socialization was that it gave the child's acquired behavioural dispositions a stability that showed an increasing independence of external control. Without using the term 'internalization' which clearly presents difficulties of definition and interpretation, Aronfreed's claim can be partly checked by the way children decide between a self-evaluation of rightness and wrongness as compared with reliance upon an external control or authority. Furthermore, children's attitudes about wrongdoing when no one would know could be investigated: this type of wrongdoing has been compared to the example of Gyges in Plato's *Republic*. Gyges, when given a ring which made him invisible, believed he had complete freedom of action, and acted wickedly.

(c) *Influences on the acquisition of moral concepts*

(1) *Home*. Various influences on children's acquisition of moral concepts have been studied. Hoffman (1963) concluded that the influence of fathers and mothers on children's moral development suggested their affecting different aspects of this process. Thus, with the father often in the background, the mother primarily determined the ways in which their children came to grips with moral issues, unless special circumstances, e.g. relating to delinquency, warranted the intervention of the father. Andry (1960) confirmed the importance of the relationship between delinquents and their fathers. On the other hand, Garai (1966) commented on the unimportant roles many fathers played in our society, being frequently absent from home; and Crane (1958) also believed that many of the boys in our culture saw their father in only a few aspects of his adult male role. Earlier Hartshorne and May (1928-30) reported a closer relation between the ideas of the child and mother rather than father, while Lee (1965), in her research study, cited results to show how the mother's influence was evident in every sphere of home life.

(2) *Peer-group*. Is the influence of parents, on moral issues, less as children become adolescents, and conversely does the peer-group influence increase? Remmers and Radler (1957) reported there was empirical evidence that most adolescents favour parental over peer advice. Morris (1958) concluded that the judgements of the adolescents in his study did not merely reflect the 'voice of the group'. These viewpoints contrast with Piaget's (1932) assertion that: 'Solidarity between equals appears once more as a source of a whole set of complementary and coherent moral ideas which characterize the rational mentality.'

(3) *Church and mass media*. What are the views of children on such influences as the church and the mass media? Is the influence of the church seen as declining but that of the mass media as extending its scope? Or, with 'normal' children, does the mass media have less influence than commonly believed? (see Searle, 1971).

Predictions

The following hypotheses were put forward for investigation:

1. Children's definitions of moral concepts would increase with age, with girls gaining higher scores than boys, and children of 'higher' social classes gaining higher scores.

2. Children's use of such terms as 'right' and 'wrong' would indicate a decreasing degree of reliance upon external authority, with older children resisting wrongdoing 'when no one would know' more than younger children. Girls would again obtain higher scores on these aspects of dealing with moral issues.

3. Influences on the acquisition of moral concepts would reveal whom the children believed helped them most when they were perplexed. It would be expected that mother's influence would be much greater with younger children, and the peer-group with adolescents; the influence of the mass media would be marked while that of the church would have declined.

Tests and Sample Used

(a) Tests used

(1) *Concepts*. Eight terms were chosen which were thought central to moral discourse from a study of the following: Graham (1972), Peters (1969), Wilson *et al.* (1967) and Wright (1971). They were 'honesty', 'responsibility', 'doing right', 'being good', 'being bad', 'guilt', 'conscience' and 'having a character'. The sample were invited to define these terms. Their replies were compared, as a criterion, to the Oxford Dictionary definitions, and marks granted on the basis of one mark for a partly grasped and two marks for a more fully grasped notion of the concept. Six experienced educational psychologists judged these replies and a close measure of agreement was achieved. Examples of the children's replies will be given in the section dealing with the results to illustrate the method by which marks were apportioned (see p. 87).

(2) *Use of terms*. Questions: (a) 'How do you know whether something is "right" or "wrong" to do?' This was an open-ended test; (b) 'Would it be right for you to do something "wrong" if no one else would know about it?' A 'yes' or 'no' response was required here.

(3) *Influences*. Questions: (a) 'To whom would you turn for help if you were not sure whether something was "right" or "wrong": yourself, friends, mother, father, relatives, teachers, or any other person?'; (b) 'Do you think the following help you to learn about "doing right", "being honest and good": films, television, newspapers, and the church?'. 'Which helps you most?'.

(4) *Intelligence tests*. Tomlinson Junior School test (1953): younger children. Richardson's Simplex Test (1932): older children.

These are both group tests, the Tomlinson being administered orally and the Simplex being a pencil-and-paper type test. Both tests embody a variety of verbal items and are essentially concerned with verbal comprehension. Except for the intelligence tests, these were set out in a questionnaire form with obviously questions 2(b), 3(a) and (b) requiring forced-choice type responses. The children were asked to underline what they believed to be their response, and in 3(b) to draw a circle around the influence which they believed helped most. All the testing was done in the respective schools with no members of staff being present; the children were told that no one in their school would see their replies, and that they had no bearing on any school work or examinations in any way, but that we were interested in their opinions on certain topics.

(b) Sample

Five groups of children from ages seven to 15 were selected by a process of 'cluster sampling' in which schools were selected at random until the numbers in the various age-groups were obtained (Butcher, 1966). The total number of children was 700, with 200 children from age-groups 11, 13 and 15; and 50 children from age-groups seven and nine—half boys and girls in each age-group. The smaller number of younger children was chosen because it was felt they could be interviewed individually and their responses recorded on the questionnaire sheet, whereas older children could take the tests in small groups. It was found necessary to use seven primary and four comprehensive schools.

The mean ages of the age-groups were 7-0, 8-10, 10-9, 12-9, and 14-11; and the mean intelligence quotients were 103·4, 103·8, 101·8, 100·3, and 99·8 respectively, with a range of quotients from 83 to 125. In respect of social classes, the sample showed close agreement with the proportions of the National Census statistics.

Table 1: Development of individual terms: percentages of 'fully grasped' replies on year groups

	7			*9*			*11*			*13*			*15*			
Years	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	
Honesty							4	7	5·5	12**	43	27·5	35*	49	42	3
Responsibility							2	5	3·5	3**	13	8	18*	31	24·5	4
Doing Right								1	·5		1	·5	1	5	3	8
Being Good										1	2	1·5	8	16	12	6
Being Bad							1	1	1	2	2	2	3*	10	6·5	7
Guilt	4	4	8	20	32	26	24	36	30	31**	53	42	62**	94	78	1
Conscience				4		4	5	4	4·5	9**	23	16	51	54	52·5	2
Having a Character								1	·5	3	9	6	14**	30	22	5

Honesty: Years 11 v 13, 15**
13 v 15**

Responsibility: Years 13 v 15**

Good: Years 13 v 15**

Bad: Years 11 v 15**
13 v 15*

Character: Years 11 v 15**
11 v 13*
13 v 15**

Conscience: Years 9 v 15**
9 v 13*
11 v 13, 15**
13 v 15**

Guilt: Years 7 v 9, 11, 13, 15**
9 v 13*
9 v 15**
11 v 13*
11 v 15**
13 v 15**

**0·01 level
*0·05 level

Statement of results

(1) *Concepts*

Age differences. The results show an increase in mean scores with age; each age-group from seven to 15 gaining statistically higher mean scores than the previous age-group. These were all significant at the 0·01 level (see Table 2).

Sex differences. Differences between the sexes show girls gaining higher mean scores on all age levels except the seven-year-old age level. At the 13- and 15-year-old levels, the girls' scores were statistically significant at the 0·01 level (see Table 2).

Social class differences. Differences between the socio-economic classes, indicated that children

Table 2: Definition of concepts: mean scores and standard deviations

Year 7	Mean Score=2·86 σ=1·65	Boys=3·12 σ=1·68	Girls=2·60 σ=1·48
Year 9	Mean Score=4·16 σ=1·65	Boys=4·08 σ=1·66	Girls=4·24 σ=1·62
Year 11	Mean Score=5·67 σ=2·13	Boys=5·44 σ=1·98	Girls=6·02 σ=2·19
Year 13	Mean Score=7·24 σ=2·28	Boys=6·25 σ=2·04	Girls=8·19** σ=2·19
Year 15	Mean Score=9·43 σ=2·28	Boys=8·88 σ=2·16	Girls=9·97** σ=2·21

** 0·01 level
* 0·05 level

Table 3: Definition of concepts: social class mean scores

	Social Class:					
	N	*1/2*	*N*	*3*	*N*	*4/5*
Year 7	12	2·2	24	2·9	14	3·3
Year 9	11	5·0*	26	4·3	13	3·1
Year 11	46	6·0	103	5·7	51	5·7
Year 13	44	7·3	97	7·3	59	7·2
Year 15	42	9·9	104	9·6	54	8·8

* 0·05 level

from classes one and two obtained higher mean scores, on all age levels except the seven-year-old age level, than children from the lower classes. At the nine-year-old level, the difference between the mean scores of classes one and two compared with four and five were significant at the 0·05 level. On other age-groups, however, the mean score differences did not reach the 0·05 level of significance (see Table 3).

Analysis of the development of individual terms. A study of the number of children on each age-group who gained 'two point' answers, or had more fully grasped the concepts, is of interest in that only two terms, namely 'guilt' and 'conscience' were successfully defined by any children aged seven and nine. Two seven-year-olds and thirteen nine-year-olds were able to define 'guilt'; and one nine-year-old defined 'conscience'. From the ages of 11 to 15, the older age-groups gained higher percentage scores for fuller definitions. These definitions were nearly all statistically significant, with the majority at 0·01 level of significance.

Girls gained a higher percentage of fuller definitions on all the terms except 'conscience' at the 11-year-old age level. Again, many of these differences were statistically significant beyond the 0·01 level.

At age 15, 'guilt'—the only term fully defined by all age-groups—was the term defined by the highest percentage, i.e. 78 per cent of the age-group (94 per cent of the girls); 'conscience' was defined by 52 per cent; 'honesty' by 42 per cent; 'responsibility' by 24 per cent and 'character' by 22 per cent (see Table 1).

Examples of how responses to concepts' test were scored:

Conscience:

2 points: 'Knowing yourself right from wrong'; 'Thinking it out whether something is right or wrong'.
1 point: 'Sense of right'; 'Sense of wrong'.
Nought: 'Think' or 'feeling' (nothing further after question).

Having a character:

2 points: 'Having a good reputation, good qualities, not weak in doing right'; 'Strong person in doing right, determined not to do wrong and not easily led to do wrong'.
1 point: 'Has qualities which are not bad'; 'Person in a play'; 'Strong morally'.
Nought: 'Being nice'; 'Get on in society'; 'Good humour'.

2(*a*) *Deciding between right and wrong*

Age differences. In deciding whether something was morally 'right' or 'wrong', the findings show that as the children grew older they claimed to decide more for themselves, and to rely less on external sources. These differences between the age-groups were generally significant at the 0·01 level (see Table 4). From 10 per cent of seven-year-olds, the percentage rose until by age 15, 78 per cent asserted they decided for themselves. Obviously, a majority of younger children cited an 'external' person to whom they turned for help. In view of difficulties in defining 'right', however, it must be questioned whether most of the sample saw the issue as a moral one, *per se*, or simply as a conflict that required a solution either by seeking 'outside' help or relying on one's own initiative.

Examples of replies to 2(*a*) by age:

7 years 'If mother says so you know it is all right.'
'Go to teacher and ask.'
9 years 'You get a little voice inside you that tells you not to do it.'
'Ask grown-ups whether it is right.'
11 years 'You stop and think if it is right or wrong.'
'By imaginary voice inside.'
'Because of 10 commandments and laws.'
13 years 'Because your conscience says to do it or not.'
15 years 'You know by thinking about it for a minute before you do anything.'
'I know when a thing is right or wrong by what I think. Your own judgment tells you.'

Sex differences. There were few differences between the replies of the sexes. At age seven boys gave a significantly higher number of responses stating they would seek the help of adults and girls of saying 'Don't know'; while at age 15, a significantly higher number of girls gave answers which indicated they were deciding for themselves (see Table 4).

Social class differences. The only social class difference was at year 11, when more children from Classes four and five said they decided for themselves (at 0·05 level of significance).

2(*b*) *'Gyges' question: doing wrong when no one would know*

There was only one significant age difference (at 0·05 level) between 13- and 15-year-olds; and, again, only one significant sex difference (at 0·01 level) at year 13 when more girls reported that such an action should not be done. However, on years seven, nine and 15 more girls believed it would be permissible if no one knew about the action, although the differences here were not significant (Table 4). Only one significant social class difference emerged, at year seven, when Social Class three obtained the highest percentage of children who believed it would be wrong to do an action even though no-one would know about the outcome.

3. *Influences*

(a) Who helps most

Age differences. Here the increasing influence of 'mother' was clearly observed; from 26 per cent at year nine, 39 per cent at years seven and 11, 48 per cent at year 13 to 57 per cent at year 15,

Table 4: Use of concepts

(2a) KNOWLEDGE OF RIGHT AND WRONG, REPLIES IN PERCENTAGES

Years:	*7*			*9*			*11*			*13*			*15*		
Method:	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*
Internal	8	12	10	16	24	20	35	34	34	40	45	43	71*	85	78
External	72**	28	50	64	60	62	44	38	41	41	40	40	24**	8	16
Don't know	20**	60	40	20	16	18	21	28	25	19	15	17	5	7	6

Years 7 v 11, 13, 15**
9 v 11*, 13, 15**
11 v 13, 15**

(2b) 'GYGES' QUESTION:

Yes	4	12	8	4	16	10	9	6	7	13	7	10	6	8	7
No	92	80	86	96	76	86	84	88	86	75**	91	83	92	90	91
Don't know	4	8	6	—	8	4	7	6	7	12	2	7	2	2	2

** 0·01 level
* 0·05 level

Table 5: Influences on acquisition of concepts: percentage replies

Years:	*7*			*9*			*11*			*13*			*15*		
	Boys	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*
Asking for help:															
Yourself	9	3	6	7	—	3	4	3	3	5	1	3	7	2	5
Friends	9	6	8	7	3	5	6	6	6	8	6	7	7	11	9
Mother	37	47	39	19	31	26	36	43	39	36**	60	48	44**	70	57
Father	36	35	38	26	38	32	33	32	33	37**	20	29	32**	11	22
Relatives	—	—	—	11	6	8	2	2	2	1	2	2	—	2	1
Teachers	6	9	8	11	16	14	12	10	11	8	6	7	3	2	2
Others	3	—	1	19	6	12	7	4	6	5	5	4	7	2	4

Relatives: Years 9 v 11, 13 and 15**

Mother: Years 7 v 15**
9 v 11*, 13 and 15**
11 v 13*, and 15**
13 v 15*

Teachers: Years 7 v 15*
9 v 15**
11 v 15**
13 v 15**

** 0·01 level
* 0·05 level

these percentages of children would first turn to their mothers for help about resolving differences concerned with right and wrong. Many of these differences between the age-group percentages were significant at the 0·01 level. 'Father's' influence correspondingly declined through the age-groups from 38 per cent at year 7 to 22 per cent at year 15. 'Relatives' also declined in influence as the children grew older; but 'friends' were named more often although non-significantly. The role given to 'teachers' varied from one age group to the next; being lower at years seven, 13 and 15, and reaching its highest point at year nine, 14 per cent (see Table 5).

Sex differences. Sex differences in the responses indicate girls naming mother more on each age-group except year nine; and boys naming father more except for years 11 and 15, but with less difference between the boys' percentages of responses for their respective parents. Few of the sex differences achieved significance except at years 13 and 15: here girls choosing 'mother' and boys 'father' were significant at the 0·01 level. Only 11 per cent of girls at 15 named 'father' compared with 44 per cent of boys naming 'mother' and 32 per cent 'father' (see Table 5).

Social class differences. None of the social class differences were significant. However, it was interesting that at year 15, 'mother' was named more by Classes four and five than by one and two; and father, in the opposite direction, being mentioned most by Classes one and two.

(b) Whether the mass media helped or not

Age differences. The replies to the question about the influence of the mass media showed many significant differences between the age-groups. For films and television, seven-year-olds gained the highest percentages, 68 per cent and 64 per cent respectively. However, although the proportions for films declined throughout the age-groups, those for television decreased at year nine and then steadily increased up to year 15 when over a half of the age-group thought television helped. Newspapers, in a similar trend, declined from year nine to 11, but then increased up to 60 per cent at year 15. There was a decline in the percentages which thought the church helped, with the lowest proportion being found at year 13. Nevertheless, a large percentage of each age-group believed the church helped them most in learning about honesty, doing good, etc. The highest figure was obtained at year seven, 92 per cent, and the lowest at year 13, 64 per cent (see Table 6).

Sex differences. There were only two sex differences which were significant (0·01 level), namely, at year 15, more girls than boys believed the church helped them most; and year 13, more boys believed they gained help from films (see Table 6).

Social class differences. Two of these differences were significant (0·05 level): at year 13, more children in Class three believed television helped, and, at year 11, more children in Classes one and two believed the church helped.

A finding to be noted is that when the development of concepts was studied, the age-groups nine and 13 at times stand apart from the others. In deciding on whether something was right or wrong, more children at year nine mentioned an external source, and the proportion of year 13 was little different from year 11. Moreover, on doing wrong when no one would know, more children (ten per cent in both cases) stated it would be permissible to do such an action at ages nine and 13. Also at nine and 13, more boys believed 'father more important than 'mother'; in comparison to more boys on the other age-groups placing 'mother' first. Pringle and Gooch (1965) have pinpointed years nine and 13 as standing out in their developmental phases of ideal person choices. It seems possible that at year 9, children are emerging from the earlier more primitive amoral phase and

Table 6: Influences on acquisition of concepts: percentage replies

Years:	*7*			*9*			*11*			*13*			*15*		
	Boys	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*	*Boys*	*Girls*	*All*
'Do these help'															
Films:															
Yes	72	64	68	44	24	34	32	24	27	50**	29	40	38	32	35
No	28	36	32	52	4	58	59	59	61	44*	61	52	57	58	58
Don't know	—	—	—	4	12	8	9	17	12	6	10	8	5	10	7
Television:															
Yes	68	60	64	40	28	34	42	37	39	58	45	52	56	51	53
No	32	36	34	60	60	60	46	47	47	36	45	40	39	41	40
Don't know	—	4	2	—	12	6	12	16	14	6	10	8	5	8	7
Newspapers:															
Yes	72	72	72	76	52	64	45	46	45	59	54	56	66	54	60
No	28	40	24	24	36	30	46	40	43	31	35	33	29	40	35
Don't know	—	8	4	—	12	6	9	14	12	10	11	11	5	6	5
Church:															
Yes	100	100	100	100	96	98	93	95	94	82	94	88	87	95	91
No	—	—	—	—	—	—	3	2	3	14	6	10	13	3	8
Don't know	—	—	—	—	4	2	4	3	3	4	—	2	—	2	1
Which helps most:															
Films	4	—	2	4	4	4	1	2	1	6	1	3	1	—	—
Television	—	—	—	4	4	4	6	2	4	9	7	8	7	3	5
Newspapers	8	4	6	16	4	10	5	10	7	5	7	6	15	9	12
Church	88	96	92	72	76	74	75	72	73	59	66	64	63**	81	72
Don't know	—	—	—	4	12	8	13	14	15	21	19	19	14	7	11

Films: Years 7 v 9, 11, 13, 15**
11 v 13**

Television: Years 7 v 9, 11**
9 v 13, 15*
11 v 13, 15**

Which helps most?: Years 7 v 9*, and 11, 13, 15**
11 v 13*

Newspapers: Years 7 v 11**, and 13*
9 v 11*
11 v 13*, and 15**

Church: Years 7 v 13, 15*
9 v 13*
11 v 13*

** 0·01 level
* 0·05 level

entering that of moral conformity. Then at year 13, a further reappraisal takes place in early adolescence, with the ability to undertake formal operational thinking, which enables the adolescent to progress towards a more autonomous outlook.

Discussion

1 *Concepts*

An analysis of the replies confirmed the hypothesis that mean scores would increase with age, with girls gaining higher scores on most of the age levels. There was a generally non-significant trend for children of social classes one and two to gain higher scores. A detailed study of the individual terms used revealed the sample having particular difficulties with 'doing right', 'being bad' and 'being good'. This was largely ascribable to the non-moral connotations of these terms, and because they were frequently equated with work in school being correct or otherwise, and obedience or disobedience to parents or teachers. Obviously, no prediction could have been made which implied that the sample would have been able to strictly define the terms used, terms which have presented perennial problems to moral philosophers. Nevertheless, there is evidence that the majority of children below year 15—to use Wilson's phrase —'do not have' a grasp of these concepts. This must be stressed as it clearly influences the way children are able to 'use' moral concepts.

Many writers, e.g. Peters (1969) and Wilson (1972) have warned against the vague and imprecise use of moral concepts. This present investigation has suggested that the misuse of such concepts by adults has an influence on children. How often, for example, do we find adults exhorting children 'to be good' or 'not to be bad'? Teachers, in the course of their duties, regularly use these terms with groups of children. Yet judging by the scores of this sample, it would appear that to many of the children such terms contained ambiguities. It is also possible that early misunderstanding might inhibit later comprehension and proper use. Thomas (1973) concluded her article, on children's use of language in the nursery, by suggesting that more attention should be given to the interaction between teacher and child, the child 'may not otherwise learn how to use speech in order to understand the continually extending horizons that are presented to him in the school situation'. Hence, there is surely a need for more detailed studies into the ways of communicating such difficult concepts to children at various stages of their development; and to assist children to use concepts relevantly in more than limited spheres, (compare Herriot, 1970, ch. 6).

There was evidence that girls gained higher scores, yet this expected finding must be interpreted cautiously. In view of the reported superiority of girls on many aspects of language development, this finding might only be confirming girls' general superiority in vocabulary and might have, therefore, less relevance to moral development *per se*. Lee (1965) suggested that girls learnt to be more articulate about their feelings and hence their story-completion endings, e.g. on guilt-type stories, were fuller and more descriptive than those of boys.

Two comparisons can be made: first Bull (1969) reported that in his study 'guilt' was typically expressed by nine-year-olds, 26 per cent at year nine, of this sample gave correct replies; while 'conscience', Bull believed, was becoming familiar at 11 and common at 13; of this sample 4·5 per cent at year 11 and 16 per cent at year 13 gave correct replies. Secondly, Kohlberg (1963) stressed that children in his sample who gained the highest scores on his test of moral judgment, tended to define 'conscience' as a choosing, self-judging function rather than as a feeling of guilt or dread. This present sample's replies at year 15 would confirm Kohlberg's view. An examination of their replies to 'conscience', showed that 78 per cent defined the term as described by Kohlberg (e.g. 'reasoning things out or using it as a guide—to choose the right thing', 'you have a choice of doing something and so you do the thing you think is right'); and 22 per cent defined conscience in terms of a feeling of guilt or dread (e.g. 'is when I feel guilty', 'you've done wrong and feel you must own up').

2(*a*) *Deciding between right and wrong*

The hypothesis was confirmed with older children increasingly stating that they decided more for themselves, and girls giving more replies indicating that there was less reliance upon external sources for help. A tendency was, however, observed of more children on years nine and 13 citing external sources. The earlier naming of parents and teachers was superseded by quasi-internal-external associations, e.g. a little voice inside telling you what to do in an authoritative manner. Later, these external influences declined and both cognitive and affective aspects were now considered along with the more positive function of knowing whether something was right to do. There were many replies indicating that emancipation from earlier authorities had not been completed. A comparison can also be made with Costanzo and Shaw's (1966) finding that conformity in their sample increased until adolescence and then decreased, with the consequent ability to have more confidence in their own judgments in spite of the disagreement of a unanimous majority.

2(*b*) *Doing wrong when no one would know*

More older children resisted wrong-doing, thereby confirming the hypothesis. Although significantly more girls than boys, at year 13, indicated it would still be wrong to commit an offence when no one would know, at years seven, nine and 15 more girls stated that it would not be wrong, the differences not being significant. Yet these results are of interest in that they went against the trend of girls gaining 'higher' scores than boys. This finding links with those obtained by Hartshorne and May (1928-30) and Terman and Miles (1963). These researches reported girls as being more likely to cheat in a party test and more dishonest on a peeping test when it was thought no one else would know. Eysenck (1965), on the other hand, refers to Drew's study on the planning of pedestrian crossings, when it was found that for women the presence or absence of a policeman had little effect on their behaviour.

3(*a*) *Who helps most*

The influence of mother is clearly apparent in this present study, being named by 57 per cent of 15-year-olds. Father's influence correspondingly declined. The prediction that the influence of the peer-group would increase at the expense of mother was not confirmed. Friends were named more often by 15-year-olds than by the other age-groups, but by only nine per cent. Hence, mother's influence confirms the findings of Hoffman (1963), Garai (1966), Crane (1958), Hartshorne and May (1928-30) and Lee (1965). That the peer-group influence was not such a major factor, confirmed the work of Remmers and Radler (1957) and Morris (1958). Both of these studies opposed the viewpoint that the adolescent was merely 'a slave to a peer-group'. Moreover, Piaget's emphasis on 'solidarity between equals' and his superficial analysis of parent/adolescent relations must be questioned on the basis of these findings.

3(*b*) *Influence of the mass media*

The present findings do not indicate the mass media having such a marked influence as was predicted. Over the age-groups, proportions of the sample stated that they were not helped by the mass media, e.g. 43 per cent at year 11 against newspapers, usually well over a half against films, and 60 per cent at year nine against television. There is thus little support to indicate an uncritical and wholesale acceptance of the mass media's influence. Brown (1963), in a summary of numerous investigations which had been made on the effects of television programmes upon children, believed that there was no proof that they had any effect on the normal child although they may have a different effect on the child who was maladjusted. Furthermore, Halloran *et al.* (1970), studying the effects of the mass media, compared the model choices of a sample of delinquents with a matched group of working class subjects and a lower middle class sample. No significant differences were found in the boys' choices; and delinquent girls chose fewer personalities from the mass media than the working-class group, but more than the middle class sample. Both Noble (1971) and Searle (1971) have found that when choosing models, mass media figures were particularly prominent in highly aggressive boys, although they pointed out that such choices did not necessarily imply any negative connotation. Thus one must be guarded in interpreting the influence of the mass media on the development of moral concepts in a sample of 'normal' school children, yet recognising that there might be differential effects in the case of maladjusted or aggressive children.

Although the incidence of people attending church is declining, substantial majorities of children on each age-group believed the church helped, and also that the church helped more than the mass media. Fewer children at year 13 believed in the efficacy of the church's help, and, in general, older children felt the church helped them less. Interpreting such figures is less straightforward, possibly lip-service is being paid to the church's influence, or that the children believed this was the response they should make, or that they were being moulded by adults to believe in the church's contribution. The problem of unravelling hypocritical poses could be taken up in future studies.

Educational implications

Educationally, it would seem that more attention should be given by teachers to their introduction of difficult concepts with, in particular, younger children. If children are to be helped 'to have' moral concepts so that they may be able to use such concepts accurately and relevantly, care must be taken with their initiation to such terms in case this impedes or distorts the later acquisitional process. The role of the teacher is, therefore, crucial. A realization must also be made of sex differences in the grasp of concepts, and support given in schools to those children who take a more critical attitude towards their peer-group and the mass media. If it is a fact that fewer people now attend church, yet substantial numbers of children believe the church helps them in acquiring moral concepts, what part should schools play? Should schools further attempt to encourage the overt recognition of the church's contribution, or perhaps more realistically, through e.g. moral education, plan to assist pupils to develop a more autonomous, authentic outlook with progressively less reliance on external authorities? Clearly, schools cannot withdraw from their involvement and the decisions made by them have a significant bearing on the way their pupils acquire, in particular moral concepts, and more generally other concepts which are of obvious relevance to their linguistic and intellectual development.

References

ANDRY, R. G. (1960). *Delinquency and Parental Pathology*. London: Methuen.

ARONFREED, J. (1969). 'The concept of internalization', in GOSLIN, D. A. (Ed.) *Handbook of Socialization Theory and Research*. Chicago:Rand McNally & Co.

BROWN, J. A. C. (1963). *Techniques of Persuasion.* Harmondsworth: Penguin Books.

BULL, N. J. (1969). *Moral Judgment from Childhood to Adolescence*. London: Routledge & Kegan Paul.

BUTCHER, H. J. (1965). *Sampling in Educational Research.* Manchester: Manchester University Press.

COSTANZO, P. R. and SHAW, M. E. (1966). 'Conformity as a function of age level', *Child Development*, **37**, 967–75.

CRANE, A. R. (1958). 'Pre-adolescent gangs and the moral development of children', *Brit. J. Educ. Psychol.*, **28**, 201–8.

EYSENCK, H. J. (1965). *Fact and Fiction in Psychology.* Harmondsworth: Penguin Books.

GARAI, J. E. (1966). 'Formation of the concept of "self" and development of sex identification', in KIDD, A. J. and RIVOIRE, J. L. (Eds.) *Perceptual Development in Children*, pp. 344–88. London: University of London Press.

GRAHAM, D. (1972). *Moral Learning and Development: Theory and Research*. London: Batsford.

HALLORAN, J. D., BROWN, R. L. and CHANEY, D. (1970). *Television and Delinquent Behaviour.* Leicester: Leicester University Press.

HARTSHORNE, H. and MAY, M. A. (1928–30). *Studies in the Nature of Character:* 3 Vols. New York: Macmillan.

HERRIOT, P. (1970). *An Introduction to the Psychology of Language*. London: Methuen.

HOFFMAN, M. L. (1963). 'Child rearing practices and moral development: generalizations from empirical research', *Child Development*, **34**, 295–318.

KOHLBERG, L. (1963). 'Moral development and identification', in STEVENSON, H. W. *et al.* (Eds.) *Child Psychology*, pp. 277–332. Chicago, Illinois: University of Chicago Press.

LEE, M. (1965). 'Children's reactions to moral transgression', Unpub. MEd thesis, University of Durham.

MCCARTHY, D. A. (1956). 'Language development in children', in CARMICHAEL, L. (Ed.) *Manual of Child Psychology*. New York: Wiley.

MACCOBY, E. E. (Ed.) (1967). *The Development of Sex Differences*. London: Tavistock Publications.

MORRIS, J. F. (1958). 'The development of adolescent value-judgements', *Brit. J. Educ. Psychol.*, **28**, 1–14.

NOBLE, G. (1971). 'Some comments on the nature of delinquents' identification with television heroes, fathers and best friends', *Brit. J. Soc. Clin. Psychol.*, **10**, 172–81.

PETERS, R. (1969). 'Moral education—tradition or reason', in MACEY, C. (Ed.) *Let's Teach Them Right.* London: Pemberton Books.

PIAGET, J. (1932). *The Moral Judgement of the Child.* London: Routledge & Kegan Paul.

PRINGLE, M. L. K. and GOOCH, S. (1965). 'Chosen ideal person, personality development and progress in school subjects', *Human Development*, **8**, 161–80.

REMMERS, H. H. and RADLER, D. H. (1957). *The American Teenager*. New York: Bobbs-Merrill.

SEARLE, A. (1971). 'A study of "admired people" among adolescents in relation to aggression and sex differences', *J. of Moral Education*, **1**, 61–6.

TERMAN, L. M. and MILES, C. C. (1936). *Sex and Personality*. New York: McGraw-Hill.

TERMAN, L. M. and TYLER, L. E. (1956). 'Psychological sex differences', in CARMICHAEL, L. (Ed.) *Manual of Child Psychology*. New York: Wiley.

THOMAS, V. (1973). 'Children's use of language in the nursery', *Educ. Res.*, **15**, 3, 209-16.

WILSON, J. *et al.* (1967). *Introduction to Moral Education.* Harmondsworth: Penguin Books.

WILSON, J. (1972). 'Problems of description in Piagetian research', *Bull Br. Psychol. Soc.*, **25**, 305-7.

WRIGHT, D. (1971). *The Psychology of Moral Behaviour.* Harmondsworth: Penguin Books.

Religious Beliefs and Moral Values

the Influence of the School

J. Marvell, St. Osyth's College of Education, Clacton

Summary

The influence of teachers on the religious beliefs and moral values of pupils in two multi-racial comprehensive schools was studied. Nearly all the teachers had come from homes where the traditional religious beliefs and moral values had been affirmed. Only one half of them, in each school, still supported a religious view of life and just over half still accepted the Christian ethic as their basis for morality. The majority, however, saw character development in one form or another as their primary aim in teaching and counted religious and moral education as part of their task. Most of the teachers wished religious education to continue in the schools but approximately half of the sample thought that it should be more broadly based than at present. The head teachers held rather conservative views. In conclusion, evidence is cited suggesting that the teachers' influence was not substantial and the view put forward that change to a more open approach in this area of school life is needed.

The decline in active participation in public religious activity and the increasing secularization of education during the last century has made it necessary to re-appraise the school's part in the religious and moral development of the child. In areas containing immigrant population the need is even more urgent. The material discussed in this paper is part of a larger study into the formative influences of home, school and community on the religious beliefs and moral values of a multi-racial group of young people.[1]

Procedure

The sample was primarily teachers at two schools situated in the same suburb of a large town in southern England and subsequently a sample of 60 children from each school, consisting of 20 English, 10 West Indian, 10 Greek Cypriot, 10 Turkish Cypriot and 10 Indian children. The schools were both mixed comprehensive, each with approximately a thousand pupils aged between eleven and eighteen. The official percentage of immigrants for one school, (A), was 34·2 per cent and the other school, (B), 42·0 per cent, although the numbers of pupils whose parents, or who themselves, had been born abroad was higher in both schools.

The methods used in studying the schools were participant observation and interviews with the staff. These were carried out during a series of regular visits for complete days over a period of five terms. School and house assemblies, lessons and extra-curricular activities were observed. In addition, every teacher who had been at the schools for a minimum of two academic years was interviewed. This amounted to 42 members out of 56 staff in school (A), of whom two were immigrants, and 60 out of 66 in school (B), of whom three were immigrants, a total sample of 102 teachers. The interviews were semi-structured, the respondents were encouraged to talk as freely and fully as possible. The length of interview varied from fifteen minutes to an hour. Those interviewed were asked to describe:

1. their religious upbringing;
2. present religious beliefs (or philosophy of life) and ethical views;
3. aims in teaching, and whether these included religious beliefs and moral values;

---

[1] Further information in a thesis for Leicester University is available from the author.

4. views on the future of school assembly, religious and moral education.

The exigencies of the school situation made the interviewing difficult on occasions and there were cases where complete information was unobtainable.

Upbringing and present religious views

Table 1 gives a picture of the information gathered on Question one and the first part of Question two.

Table 1 indicates that over 90 per cent of the teachers came from backgrounds where the traditional religious beliefs and moral values were affirmed. In over 40 per cent of the homes there was active participation by one or both of the parents in public religious activity. As over 90 per cent were sent or taken to Sunday school or church, it is probably safe to assume that those home backgrounds designated as 'not ascertained', were, in the main, nominally religious. If this is the case, the figure has some correspondence with national figures (Argyle, 1958; Martin, 1967; Wilson, 1969).

The figures, in Table 1, for the teachers' present views indicate a division of opinion. Approximately half of the staff in each school affirm and half deny or question the traditional beliefs of our society. But in a study of some of the homes of pupils in these schools it was found that the majority of parents nominally affirm traditional religious beliefs, a point confirmed for English parents in May and Johnston's study (1967).

Table 1: Teachers' religious upbringing and present views on religion

Home Influence	% School A	% School B	% A+B
Practising	45·0	42·8	44·1
Nominal	31·6	30·9	31·3
Non-religious	10·0	9·5	9·8
Not ascertained	13·3	16·6	14·7
Church Influence			
Attended Church/ Sunday school	96·6	85·7	92·1
Present View			
Practising	28·3	26·1	27·4
Nominal	23·3	16·6	20·6
Agnostic	28·3	21·4	25·5
Atheist	20·0	35·7	26·5

Table 1 also indicates that a quarter of the staff at both schools had a strong practising commitment towards religion. Within this small group, approximately a half (54·5 per cent School A, 47·0 per cent School B) were practising members of the Church of England, a group which one would suppose was even more strongly committed to upholding the traditional beliefs and values of our society.

Ethical views

When questioned about their present ethical views over half the teachers (63·3 per cent School A, 52·2 per cent School B) stated that they still accepted the Christian ethic as their basis for morality. The respondent who said:

'. . . basically the teaching of Christ with the ten commandments' and the one who declared:

'I feel strongly that the Christian morality is the basis of life', represent this group's point of view.

There was considerable division of opinion, among the remaining teachers in both schools. It would seem that once the traditional religious beliefs of society are abandoned, there is a tendency for the social and ethical systems based on it to be questioned as well. Even some of those teachers wishing to affirm Christian morality demonstrated this tendency with such qualifications as:

'. . . but I feel tolerance of the opposite view to the traditional Christian ethic is essential', and

'I keep to the Christian ethic except on sex. I no longer feel chastity in and out of marriage is essential.'

Some teachers expressed more radical views:

'There are no absolutes in morals, they are relative to the situation', and

'. . . every man must pursue his own interest.'

Bases for morality suggested by those not subscribing to a Christian view included human dignity, moral relativity, Communism, conscience and personal inclination.

Teaching aims

Staff gave a variety of views when discussing their aims in teaching, as set out in Table 2. If more than one aim was mentioned, that to which highest priority was given by the teacher was used for purposes of analysis.

Table 2: The teaching aims of the staff

	%	%	%
Communal Aims	School A	School B	A+B
Citizenship	20·0	26·1	22·5
Home life	5·0	2·3	3·9
Christianity/ Christian values	18·3	14·2	16·6
Moral values	5·0	2·3	3·9
TOTAL	48·3	44·9	46·9
Individual Aims			
Pastoral care	3·3	7·1	4·9
Independent thought	8·3	4·7	6·8
Develop character	15·0	7·1	11·7
Learn subject skill	13·3	14·2	13·7
Children's enjoyment	1·6	4·7	2·9
TOTAL	41·5	37·8	40·0
Teacher uncertain	5·0	7·1	5·8
Not ascertained	5·0	9·5	6·8
TOTAL	10·0	16·6	12·6

Table 2 indicates that the teachers in both schools were divided in their aims. The division appears to be between those who saw education as a process of socialization and those who saw it in terms of benefit for the individual. Only a small minority thought that the imparting of subject skills was the principal part of their task. The majority saw it in terms of norms and values. The following statements were typical of the different categories:

Communal aims:

'The development of the child's character for society.'

'. . . to get children to see the importance of old-fashioned virtues.'

'Certain standards have to be kept . . . it is important to put over to all and sundry consideration for others.'

Individual development:

'. . . children should be given the tools of thought which will save them from being duped.'

'Enabling children to enjoy themselves and to develop their character.'

'To enable less able girls to discover their abilities.'

Subject skill:

'Developing skill in their hands.'

'. . . teach children to love music.'

Seventy-three per cent of the teachers stated that they accepted religious belief and moral education as part of their task. Further analysis, as shown in Table 3, indicates that the way that they approached this part of their work varied.

Aims in moral and religious education

The teachers' views on moral and religious education are set out in Table 3. When the answers of those who stated that this was part of their task were analysed a divergence of teaching styles became apparent.

The tendency towards authoritarianism in School A and democracy and permissiveness in School B reflects not only staff views but those of the head teachers. The divergence was apparent in many ways in the life of the schools. In School A the wearing of school uniform was rigorously enforced; silence was maintained before and after assembly; pupils stood when staff entered a room; movement around the school was strictly regulated. By contrast, in School B, uniform was little in evidence; children chattered before and after assembly; there were no rules governing movement around the school and children were not obliged to stand when a member of staff entered the room. Typical comments, illustrating the differing teaching styles were:

Authoritarian:

'I put forward a Christian view.'

'Standards must be imposed.'

Democratic:

'I feel that all sides should be considered; I am very much against dogmatism.'

'You can have discussion, and to some extent advice. . . I say what I would do myself.'

Table 3: Teachers' aims with regard to moral and religious education

View	% School A	% School B
1. Agree RE/ME part of aim	76·6	68·8
(a) Authoritarian	43·3	26·1
(b) Democratic	28·3	28·5
(c) Permissive	5·0	14·2
2. Disagree RE/ME part of aim	23·2	30·2

Permissive:

> '. . . I am wary about putting forward any religious or moral point of view.'
>
> '. . . I seek to avoid it, but give neutral answers when questioned by the children.'

It is clear that although the majority of teachers in the sample considered the matter of religious beliefs and moral values was an important part of the school's task, they were far from being convinced that the dogmatic assertion of traditional beliefs and values was correct. The relevant paragraphs of the 1944 Education Act, at the very least, enjoins the school to strongly commend them.

The future of religious education: teachers' views

Table 4 shows the teachers' reactions to this and their views as to the future of religious education and school assembly.

Table 4: Teachers' views on the future of religious education and assembly

Religious Education	% School A	% School B	% A+B
As now	26·6	30·9	28·4
Broader basis and approach	55·0	40·4	49·0
Moral education only	1·6	4·7	2·9
Abolish	3·3	4·7	3·9
Uncertain	3·3	4·7	3·9
Not ascertained	10·0	14·2	11·7
Assembly			
As now	30·0	42·8	35·5
Broader basis and approach	23·3	9·5	17·6
Non-religious			
(moral)	20·0	16·6	18·6
(secular)	1·6	4·7	2·9
Abolish	15·0	11·9	13·7
Uncertain	—	4·7	1·9
Not ascertained	10·0	9·5	9·8

These figures indicate that the majority of the teachers wanted this area of school life retained. One third appeared content with the situation as it was, implementing the 1944 Education Act. About a third thought that the school should provide a choice in matters of faith, while a minority wished to exclude it. Examples of teachers' comments illustrating these opinions were:

To keep it as now:

> 'I would keep RE and assembly much as at present.'
>
> 'RE and assembly are vital to school life.'
>
> 'RE should stay as it is now, comparative religion would bring doubts.'

To put it on a broader basis:

> 'I would abandon assembly but keep religious education on a broad basis making it voluntary after the third form.'
>
> 'I would keep RE as an option in the secondary school with a wide choice offered. I would like to see assembly develop on moral uplift lines.'

In favour of abolition:

> 'I believe religious education in its present form is outdated and even harmful in its effects.'
>
> 'I am anti school RE especially assembly.'

In both schools the religious education departments reflected the conservative outlook of the persons in charge of them (c.f. those cited in Loukes, 1965). Their stated aims were:

> 'To make Christians . . . it is important that they should be carefully guided along the right channels of thought.'
>
> 'To present the gospel and to show the way in which Christianity offers the right relationship between God and man, and to show the standard of social ethics inculcated in the teaching of Jesus.'

The RE departments instead of being central to the schools' task of socializing the pupils into the norms and values of society, as intended by the 1944 Education Act, stood at one end of the spectrum of staff aims and views. The departments, committed to the erection and maintenance of the traditional structure of society, were unable to do so because the vital consensus was lacking. They were no longer able to 'nurture Britannia's children in the family faith' (Alves, 1968).

The head teachers' attitudes

The head teachers of both schools were interviewed at the beginning and end of the study. In addition to this, opportunities arose from time to time for informal conversations. The head teacher from School A came from a devout Jewish home. She still retained a belief in God, but was only nominally religious in practice, and somewhat against religious institutions. She stated that she believed religion should be closely linked to life in the school situation. She was strongly authoritarian in the matter of morality and thought that religious belief was essential to its maintenance. This was reflected in the school, particularly in the way in which she conducted assembly. When questioned about the future, she said that she would retain assembly in her school as long as she had the right to do so. She also commented that she would provide voluntary religious education, even if compulsory religious education were abolished.

The head teacher from School B came from a devout Methodist home. He was still a very active lay preacher and member of the Methodist church. His view was that Christianity should provide the basic values of school life and that religious education and assembly be retained within a Christian framework. This was reflected in his assemblies which were usually centred on an exposition of a passage from the Bible. He expressed strong views on democracy, putting great emphasis on the school council. When on one occasion the school murmured their disapproval in assembly of his action, the head told the school that the matter would be decided by vote at the next assembly. From the foregoing, it can be seen that although the staff, in each school, to some extent reflected the view of the head teachers on authority in the school, they did not reflect their religious viewpoints.

Who influences the pupils?

A later inquiry with the sample of children from the two schools revealed that their teachers were a minor element of influence in their lives compared with the family. It became clear, however, that alternative points of view were first encountered at school. In particular, observation of lessons and other school activities demonstrated the many occasions when matters concerning religion and morality are discussed between pupils and teachers. One instance occurred during a science lesson when the question of death arose out of a discussion of pollution. The class asked the teacher, who was a Catholic, for his view and he explained it. Another occasion arose during a needlework lesson when relationships within the home were discussed. Yet a further illustration was the discussion that came out of the moral implications of a practical mathematics problem with which the class was engaged. In addition, many personal interactions between pupils and teachers conveying, through implication, the teacher's beliefs and values were observed.

Several researches in recent years have shown that the religious provisions of the 1944 Education Act have not worked as intended (Alves, 1968; Daines, 1965; Hyde, 1965; Loukes, 1961 and 1965). The changes in society are, as Bernstein (1971) argues, affecting the school, in this as in other areas of its life. As Berger (1967) points out, the divine world, or 'sacred canopy' as he calls it, is no longer believed in and accepted as sanctioning the society's structure of norms and values. Luckmann (1967) suggests, '. . . with the pervasiveness of the consumer orientation and the sense of autonomy, the individual is more likely to confront culture and the sacred cosmos as a "buyer". Once religion is defined as a "private affair" the individual may choose from the assortment of "ultimate meanings" as he sees fit . . . .' The dilemma is even greater in multi-racial schools where the children come from a variety of backgrounds, in this case, there is neither a consensus of school staff nor parents as to which religious and moral norms could unite the school as a community. Perhaps the answer lies in considering the school of the future as an association of people through which the individual pupil is given assistance in making rational and autonomous choices in the areas of religion and morality rather than as a community in which he is socialized into a set of traditional beliefs and values.

References

ALVES, C. (1968). *Religion and the Secondary School*, 213. London: SCM Press.

ARGYLE, M. (1958). *Religious Behaviour*, 23–7. London: Routledge & Kegan Paul.

BERGER, P. (1967). *The Sacred Canopy*. New York: Doubleday.

BERNSTEIN, B. (1971). 'Open schools, open society', in COSIN, B. R. *et al.*, *School and Society*. London: Routledge & Kegan Paul.

DAINES, J. W. (1965). 'An inquiry into the methods and effects of religious education in sixth form boys'. University of Leeds, Institute of Education.

HYDE, K. (1965). 'Religious learning in adolescence'. University of Birmingham, Institute of Education.

LOUKES, H. (1961). *Teenage Religion*. London: SCM Press.

LOUKES, H. (1965). *New Ground in Christian Education*. London: SCM Press.

LUCKMANN, T. (1967). *The Invisible Religion*. New York: Macmillan.

MARTIN, D. (1967). *A Sociology of English Religion*, 34–51. London: Heinemann Educational Books.

MAY, P. R. and JOHNSTON, O. R. (1967). 'Parental attitudes to religious education in state schools', *Durham Res. Rev.*, 18, 127–39.

WILSON, B. R. (1969). *Religion in Secular Society*, 21–39. Harmondsworth: Penguin.

Correction

Due to an oversight, certain incorrect figures appeared in the article 'Authoritarian personality traits among students' by J. Dunham in the last issue of *Educational Research* (16.1, November 1973). In the first full paragraph on page 41 there is a reference to 52 first-year students, 26 of whom were in an architecture class, 11 in an engineering class and 15 reading sociology; towards the end of the paragraph it is reported that 14 per cent of the sociologists were authoritarian and 35 per cent non-authoritarian, and 46 per cent of the architects were authoritarian and 20 per cent non-authoritarian. We are now informed by the author that these figures should have been as follows: 52 first-year students, 15 in an architecture class, 11 in engineering and 26 in sociology; 4 per cent of the sociologists were authoritarian, 38 per cent non-authoritarian, and 33 per cent of the architects were authoritarian and 20 per cent non-authoritarian. We apologize for this error.

Note to readers

The editors would like to make known to readers their regret for the fact that the article 'Some determinants of teacher morale in Australia' by G. M. Coverdale in our last issue (16.1, November 1973) had already been published in *Educational Review* shortly before the appearance of our own journal. Everything possible is done to avoid such occurrences, but responsibility for them ultimately lies outside the control of either journal, and the regret is shared by the editors of both. Readers have a right to expect that each issue of a journal will contain original material, except where an article of exceptional merit may have been republished because it is no longer easily accessible, in which case this will have been clearly stated. Potential contributors, too, have a right to expect that journal space will be used to fullest effect, to the best of the editors' judgement. We trust readers will appreciate that in the last resort editors have to rely on the good faith of contributors.

Religious Socialization in the Junior School

Peter Jarvis, Dudley College of Education

Summary

Most research into religious education has concentrated upon older children, but this paper reports findings obtained from the teachers of junior (seven to 11 years old) school children. From the content of religious education lessons and the overall religious sub-cultures of classroom and school, conclusions are drawn about the religious socialization process of the child. A distinction is made between the child's own religious quest for meaning and the Christian culture which is generally accepted in Britain and taught in school. The child's religion is divided into the cognitive and the emotive while the socialization process is treated as cognitive, explicit and implicit. The general conclusion reached is that religious education lessons do not respond to the child's own cognitive level and are thus likely to confuse, whereas the religious life of classroom and school leads to an emotive affinity with Christianity. Thus ambivalence towards the Christian religion is produced in the child.

Introduction

Among the functions that religion fulfils is that of providing a meaning system to the universe. It answers the unanswerable! Life is conducted in an ever expanding universe; for the small child this is bounded by family and playgroup, then the horizons broaden as he goes to school and interacts with a wider circle of peers. At most levels there are questions of meaning that demand either a cognitive or emotive answer. Initially such answers are discovered in the parents, for, as Piaget says, 'the child spontaneously attributes to his parents the perfections and attributes which he will later transfer to God . . .' (1929, p. 354). At a later stage neither the family nor knowledge gained through empirical evidence satisfy. However, questions of meaning still demand an answer, for religious questions are an 'anthropological necessity' (Luckmann, 1970).

As the child's experience broadens his questions of meaning range more widely. Piaget described a fourfold typology through which the child's religion develops. This paper accepts his basic thesis that there is an ever increasing expanding individual quest for meaning which is religious and universal.

In any era widely accepted answers to such questions form a 'theology', or a cosmology, of the social group. These answers permeate its culture and religion thus becomes a cultural system (Geertz, 1968). In Britain Christianity has been that religious cultural factor for well over a thousand years. Some still hold that Britain is a Christian country (May and Johnson, 1968). It would perhaps be more precise to suggest that today British culture contains many norms and values which are Christian in origin. Recent surveys have indicated that Christianity is still the religious reference for many people in Britain, despite the decline in church attendance (ITA, 1970).

Hence it is into this cultural system that children are socialized. Among the many agencies of religious socialization the school has become the most important in many instances. The 1944 Education Act specifies that each school day should begin with an act of worship and that regular religious instruction should be given to the pupils. Although the Christian faith is not

---

Acknowledgements

The author wishes to thank C. R. Hinings who supervised this research and read through this draft. Also E. C. Prichard who read through the draft and made a number of helpful suggestions. Warley Education Authority were generous in their assistance while the research was in progress.

specified, this is generally assumed. Christianity also appears in other elements of the school curriculum: art and craft, music, etc. There are also other religious events (Christian) in the school calendar, e.g. harvest festival and the nativity play.

The child, then, is born into a society whose culture contains much which is Christian in origin. At an age when he has few reference groups the child becomes a member of a social group whose culture is even more strictly orthodox Christian. It is likely that he will be socialized into some form of acceptance of the Christian religion before he is old enough to understand its complexities.

Broadly, socialization is postulated as a tri-dimensional process in the school group. Cognitive socialization is the actual learning from the content of the curriculum. However, in the research reported here only the content of the religious education lessons were fully investigated rather than the religious/Christian content of the whole curriculum. Explicit socialization is that process whereby the child is taught to conform to the norms and values of the group. This occurs in response to the demands of a significant other, e.g. the class teacher. Implicit socialization is the acquisition of beliefs, norms, values and traditions through imitation and experimentation. These three processes of socialization co-exist and intertwine, their separation here being artificial and employed solely for analytical purposes.

There are then two parallel processes: the growth of religious questioning in the child which does not have the presuppositions of any established religious system such as Christianity; and the internalization of the culturally accepted system. This paper examines the latter aspect of this process while noting the need for much more extensive research into the former.

Background

The actual content of the religious education and other lessons is left very much to the class teacher, and as David Martin (1967, pp. 88f.) has said, many primary school teachers are well disposed towards Christianity. In this investigation 24·79 per cent of the sample who responded affirmed that they believed in the traditional Christian doctrines and attended church at least once per month, while 72·59 per cent of the respondents described themselves as Christians. The socializing agent is the teacher. The role he plays in the three dimensions of socialization is examined, with some emphasis being placed on the cognitive one. If the class teacher is to be an effective agent in religious socialization he should be aware not only of the level of the religious maturity of the children he is teaching in order to respond to their own religious questions but also know whether the content of his own religious education lessons can be properly understood by the class.

It was anticipated that few teachers would consider the development of the child's own religion, and this assumption was shown to be correct. Hence there was considerable concentration in the research upon the latter aspect of his role. A method was devised to assess the teacher's awareness of the conceptual level necessary to comprehend the content of his religious education lesson. It was based on Piaget's findings that children of junior school age (seven to 11 years) conceptualize in concrete terms. Only at a later age is a young person able to conceptualize in abstract form. While Kohlberg (1967, pp. 85-7) has modified Piaget's scheme, this in no way negates the approach employed here.

In this paper, two basic hypotheses of the research are discussed. The initial hypothesis is that much of the level of the material employed in religious education is beyond the level of the conceptual development of the children. It is suggested that this will fail to answer the child's own religious questioning and that he will reinterpret the religious material he receives within his own conceptual framework. This will lead to either ambivalence or rejection at a later level of development. Secondly, the explicit and implicit processes of socialization are an important factor in the primary school, which may lead to an affective affinity with Christianity in later years. However, this will only increase the ambivalence that is later felt towards Christianity, since the culturally accepted religion has failed to answer the young person's questions of meaning.

The method employed in this research was a questionnaire survey of an entire population. It was conducted in one local education authority area and every junior school teacher was asked to complete the questionnaire on the same day towards the end of the spring term 1971. There were 238 replies, a response rate of 74·4 per cent if the head teachers are included within the total population although they were not asked to complete the questionnaire. Not all of the respondents answered every question.

This LEA have a small working party devising a new agreed syllabus for the junior schools and the group helped with the preparation of the questionnaire and the arrangements for the survey. Among the questions asked were those which would indicate the teachers' role in each of the dimensions of socialization and those which showed their awareness of the level of conceptual development necessary to understand the content of their lessons. The remainder of this essay reports the findings of this research and falls into two main parts, each discussing the data relevant to one of the hypotheses.

Cognitive socialization

To understand this element it was necessary to investigate the content of the religious education lessons received by the children. However, in itself this information is insufficient since cognition implies that children comprehend the lesson. Thus it was necessary to evaluate the teachers' awareness of the conceptual level necessary if the children were to understand the content of the lessons. It was expected that the lesson content would be almost completely about the Bible and Christian religion and these were the subjects focused upon in this aspect of the investigation. Before discussing this in detail it is necessary to demonstrate that the expectation was justified.

From the replies received it can be safely assumed that the vast majority of children received at least one religious education lesson per week, and many considerably more. 22·69 per cent of the respondents said that they conducted two religious education lessons weekly and 6·30 per cent had three or more. Of the teachers who had separate lessons (145), as opposed to an integrated approach, an average of 1·6 lessons a week were conducted. 11·34 per cent of the respondents did not teach the subject but in a number of instances a 'reserve teacher' took that lesson.

The actual aims that the teachers had for the lessons have been discussed more fully elsewhere (Jarvis, 1972a). The communication of moral ideas, a knowledge of Christ's teaching and of the Bible were the most frequent. The prevalence of moral education within religious education was of considerable interest and has been discussed in another article (Jarvis, 1972b). Table 1 summarizes replies to an open-ended question asking class teachers their general aims in religious education.

Table 1: The general aim of the religious education lesson

Response	*Respondents* *No.*	*Respondents* %
Teach about God	41	17·23
Teach about Christ	40	16·81
Teach Christianity	34	14·29
Prepare children to become Christians	22	9·24
Teach Bible knowledge	41	17·23
Show cultural aspects of religion	9	3·78
Social/Moral training	104	43·72
Deepen awareness of life	35	14·71
Let children ask religious questions	3	1·26
Fulfil duty, etc.	7	2·94

Some of the above elements are not explicitly cognitive and will be omitted from the present aspect of discussion. The prevalence of Christianity and Biblical knowledge is clear. Few teachers report aims that reflect any awareness of the children's own developing need to ask religious questions. Clearly teachers seek to present children with Christian and Biblical knowledge and hope that they will learn it or enjoy it as a 'good story'. It is doubtful whether the teachers have the intention of helping the children to appreciate the religious, or philosophical, meaning underlying the story, myth or legend.

The data in Table 1 were checked in a further question which asked teachers to describe the

Table 2: The content of the last religious education lesson

Response	*Respondents* *No.*	*Respondents* %
Biblical	87	36·55
Devotional	12	5·04
Linked to Christian broadcast	5	2·10
Linked to Christian person	18	7·56
Linked to Christian season	13	5·46
Other/Historical religions	6	2·52
Linked to Nature	3	1·26
Moral topic	19	7·98
Other theme/topic	15	6·30
No reply	60	25·21
Total	238	99·98

aim, content and method of the religious education lesson last conducted prior to completing the research questionnaire. From the replies it was possible to construct a summary of the content of those lessons (see Table 2).

This confirms the conclusions already drawn. It is especially significant to see how few lessons are linked to the children's own religious experience or quest. Since children are expected to comprehend the cultural religious system it must be recognized that learning ability is related to the conceptual level required to understand the material. The teachers' own awareness of the conceptual level necessary to understand their lessons needs further investigation.

Awareness of conceptual level necessary for comprehension

Obviously all primary school teachers are not familiar with the theological works available which would enable them to obtain a correct exegesis of Biblical stories, festivals of the Christian calendar and their mythical epiphenomena. Two complementary questions were constructed which attempted to assess the teachers' awareness of the level of conceptual development necessary to comprehend an academically acceptable meaning of a subject which they would be prepared to teach. Subjects were pre-set and the respondents asked to indicate those which they would teach to their present class. Some subjects were considered suitable for infant, some for junior and other for secondary children. One question was about the Christian festival of Whitsun and the other based on the Bible. While this is not a very sophisticated method it is considered adequate for the purpose of this research, for, briefly, if the interpretation of the subject suggested requires formal operational thought for comprehension it was considered too difficult for junior school children. The West Riding Agreed Syllabus (1966), which was chosen since it has been considered among the most influential, was also used to guide the selection of topics.

In the analysis of the questions the subjects are grouped, although they were placed indiscriminately in the questionnaire. Group A are suitable for infant children, B for junior children and C for secondary and above. Respondents were awarded negative marks if they ticked subjects of either Group A or C and positive marks for each topic chosen from Group B. In

Table 3: Subjects that junior school class teachers would employ in a religious education lesson approaching Whitsun

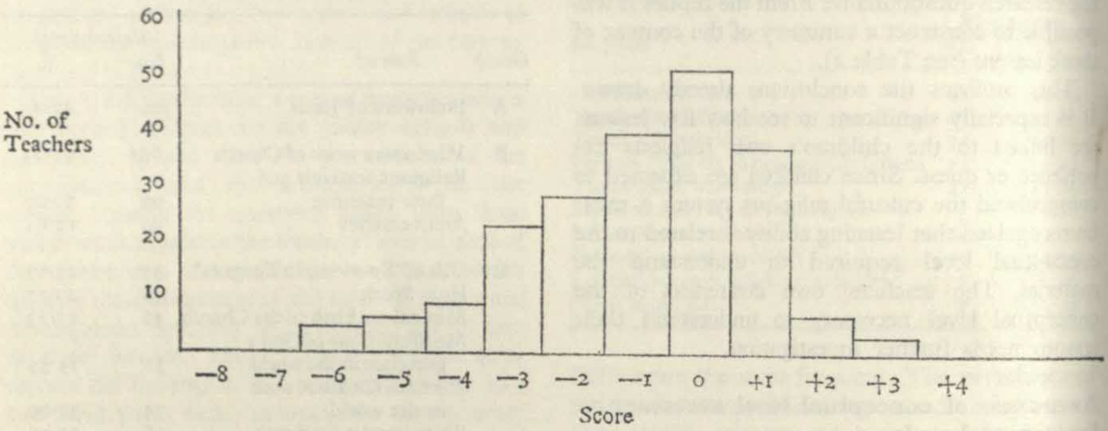
Group	*Subject*	*Respondents* *No.*	*%*
A	Followers of Jesus	101	42·44
B	Missionary work of Church	76	31·93
	Religious festivals and their meaning	93	39·07
	Anniversaries	30	12·61
C	Gift of 'Speaking in Tongues'	33	13·87
	Holy Spirit	48	20·17
	Miraculous birth of the Church	36	15·13
	Manifestations of God's presence in the world	37	15·55
	Power of Christ at work in the world	74	31·09
	Pentecostal experience	36	15·13
	No response	30	12·61

Table 4: Biblical subjects which junior school class teachers would be prepared to teach their present class

Group	*Subject*	*Respondents* *No.*	*%*
A	Andrew and Peter—fishermen friends of Jesus	172	72·27
B	David, shepherd of the sheep	156	63·55
	Christ telling Peter to forgive seventy times seven[1]	118	49·58
	Christ's encounter with Zaccheus	125	52·52
	David forgiving Saul	111	46·64
	Compassion—story of Good Samaritan	196	82·35
	Gratitude of the leper healed by Christ	172	72·27
C	Christ healing the man sick with palsy	138	57·99
	Daniel in the lion's den	115	48·32
	Christ's temptations	104	43·70
	Christ's Ascension	64	26·89
	Christ's teaching about punishment of hell for sinners	18	7·56
	Christ walking on the water to his disciples	82	34·46
	No reply	17	7·14

[1] This question may have been slightly affected by an error in the survey questionnaire which read '. . . seven times seventy'.

Figure 1: Aggregate scores of class teachers' awareness of the conceptual level of thought required to understand suggested Whitsun lesson topics



retrospect it may have been harsh to have awarded a negative score for Group A.

Taking the question about Whitsun first, the class teachers were asked, 'In a religious education lesson approaching Whitsun I would teach my class about:' There were ten pre-set replies and a space provided for any other suggestions. The replies are recorded in Table 3.

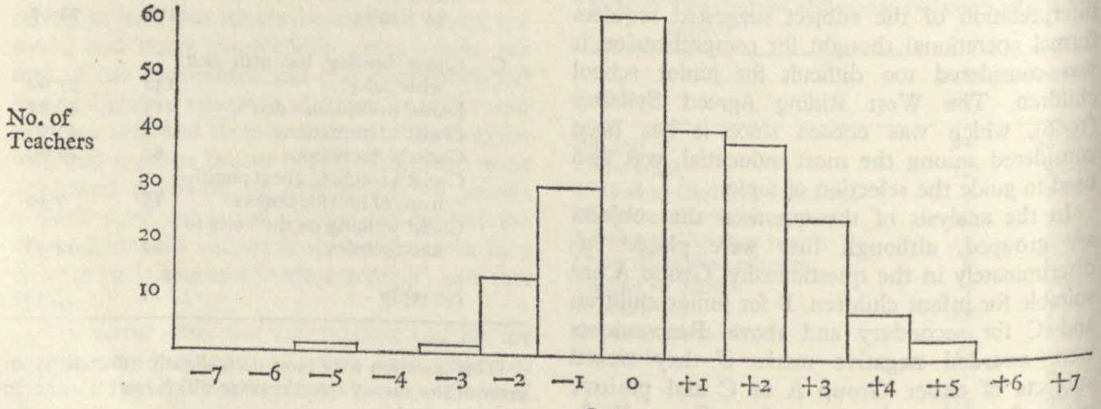
There were only six written-in answers plus eleven who said that they would not teach any specific Whitsun topic. Hence these are not enough to affect the general picture. Aggregating the scores for all the respondents a histogram (Figure 1) can be constructed.

The histogram reveals that overall 45 (18·91 per cent) of the respondents obtained a positive aggregate. The mean score is −0·75 (2 dec. pl.) which suggests that the teachers are not always aware of the level of conceptual development required to understand these topics.

There is some likelihood of a negative score in this instance because less than half of the pre-set answers are considered suitable to children in the junior school age range. Consequently the Biblical question contained a greater number of suitable subjects. The multiple choices and the teachers' replies are contained in Table 4.

There were nine respondents who ticked all 13 answers, and there were no answers in which only the six subjects in Group B were selected.

Figure 2: Aggregate scores of class teachers' awareness of the conceptual level of thought required to understand the suggested Biblical lessons



The same method of aggregating the scores is now employed as in the previous question, but the histogram in Figure 2 now has a wider spread since there are more possible answers.

In this instance 121 (50·84 per cent) of the respondents had a positive aggregate and the mean score was +0·82 (2 dec. pl.). Since there was a greater proportion of suitable answers there was more possibility of a positive score. Yet the proportion who obtained a positive aggregate was far higher than it would have been had the answers been consistent with the earlier question. This could be adduced as added support for the idea that the Bible is central to religious education in the junior school.

It is clear from these two questions that many subjects demanding a considerable degree of abstract thought in order to be understood or discussed are being taught to junior school children.

Conclusions relating to cognitive socialization

From the evidence presented above it is considered that Christian material is transmitted to the children but the transmitting agents do not demonstrate an awareness of the conceptual level necessary to understand their lesson content. It is suggested that the children will re-interpret the lesson content within their own conceptual framework, but that this will not be a correct exegesis of the religious cultural material. More importantly, most of the lessons do not occur in response to the children's own religious quest and thus appear irrelevant to them. Goldman (1964 and 1965) concluded that by the time children are capable of formal operational thought they had lost interest in religion. He, wrongly, in my opinion, tends to equate religion with Christianity and for him it may be Christianity which the children are tired of. This is supported by Loukes (1961 and 1965) who discovered that the young person was not disinterested in religion but was confused and tired of the constant 'given' of Christianity.

Our research findings both support the initial hypothesis and are additional evidence in support of these earlier works.

Emotive socialization

It is suggested that within the explicit and implicit processes emotive socialization occurs.

Table 5: The religious education lesson is designed to teach the class to:

Response	*Respondents* *No.*	*%*
Appreciate religion as a subject	22	9·24
Be religious	19	7·98
Become more moral	112	47·06
Become Christians	70	29·41
Be better citizens	98	41·18
Understand our cultural heritage	53	22·27
Appreciate events in Nature	61	25·63
No reply	24	10·08

The teachers were asked how they sought to influence the children in their religious education lessons. The multiple choices presented to the teachers in this question are indicated by the different categories in Table 5.

Since the teachers could have had more than one objective they were invited to indicate as many of the categories as they wished. Again the social-ethical responses were the highest. A comparatively large number had a confessional aim and few had a strictly educational objective.

The class teachers were also asked if they sought to communicate an emotive content which could be termed a religious feeling. Sixty-one (25·63 per cent) of the respondents affirmed that they did, a further two said that they did sometimes. Of the 61, 26 indicated that this led to an act of worship in the classroom. Many teachers said that they provided experiences for the children which would make them more aware of a 'religious' element in our humanity. Among these were both a religious response to the wonders of the world, nature and life, and an admiration of Christians who served other people.

Since 165 (69·33 per cent) of the respondents said that they had given the children no instruction in any religion other than Christianity it is safe to consider that the worship, emotion, etc. evoked would be within a Christian context. It is recognized that measuring the effects of such explicit socialization is difficult but the above data indicate a wide range of positive influences provided by the class teachers. Coupled with the implicit socialization it is felt that this will effect an effective response by the children towards Christianity.

The school religious culture is outwardly Christian. Most schools abide by the 1944 Education Act which states that every school day

Table 6: Frequency of prayers in the classroom

Frequency per day	*Respondents* *No.*	 *%*
Once	81	34·03
Twice	17	7·14
Three or more	4	1·68
Occasionally	4	1·68
None	119	50·00
No reply	13	5·46
Total	238	99·99

should begin with an act of worship. Many schools use the School Broadcasts which are usually Christian in aim and content and were so during the time of this research. The hymns sung during assembly are always Christian and many schools set aside periods in the timetable for practising these hymns. Other lessons within the curriculum have a Christian content but these were not investigated here.

Within the classroom there is also a Christian element in the sub-culture. Teachers often lead corporate Christian prayer in the classroom, and they were asked the frequency of this (see Table 6).

Most teachers who have classroom prayers do so at the end of the day, at the end of the morning, in religious education lessons, or before milk. (One of the unforeseen consequences of the government withdrawing free milk from children over seven years old was a slight lessening in the frequency of classroom prayers!) However, the significant thing is that nearly half of the teachers have classroom prayers in addition to daily school assembly. Clearly in the junior school the child is exposed to considerable emotive socialization. The child undergoes a value shaping process and acquires an affective affinity towards Christianity. He also lives in a cultural milieu in which the Christian faith is loosely accepted. Thus the national culture does not conflict with his feelings towards Christianity, but neither does it strongly reinforce them. However, this discussion demonstrates the validity of our second hypothesis.

Conclusion

We have seen that the junior school creates a situation of religious cognitive confusion but has not responded to a quest for meaning which may develop in the child in relation to his universe. At the same time it has socialized the child into some affective affinity towards the Christian faith. Hence we can begin to understand why there is frequently a residual sympathy, although rarely a rational comprehension of, the culturally accepted form of religion. Supported by emotional attachment the cultural religion is resistant to change, despite suffering a 'crisis of credibility'.

The conclusion drawn from this discussion about socialization can be stated quite briefly: the result of undergoing religious socialization in the junior school is ambivalence towards the cultural religion but little nurture is given to the child's own religious quest.

References

Geertz, C. (1968). 'Religion as a cultural system', in Banton, M. (Ed.) *Anthropological Approaches to Religion*. London: Tavistock.

Goldman, R. J. (1964). *Religious Thinking from Childhood to Adolescence*. London: Routledge & Kegan Paul.

Goldman, R. J. (1965). *Readiness for Religion*. London: Routledge & Kegan Paul.

Independent Television Authority (1970). *Religion in Britain and Northern Ireland*.

Jarvis, P. (1972a). 'Teachers report their aims', *Learning for Living*, May.

Jarvis, P. (1972b). 'Religious education as a vehicle for moral education?', *J. Moral Educ.*, **2**, 1, Oct.

Kohlberg, L. (1967). 'The development of children's orientation towards moral order: 1, sequence in moral thought', cited in Higard, E. R. and Atkinson, R. C., *Introduction to Psychology*. New York: Harcourt, Brace & World.

Loukes, H. (1961). *Teenage Religion*. London: S.C.M. Press Ltd.

Loukes, H. (1965). *New Ground in Christian Education*. London: S.C.M. Press Ltd.

Luckmann, T. (1970). *The Invisible Religion*. London: Macmillan.

Martin, D. (1967). *The Sociology of English Religion*. London: Heinemann.

May, P. R. and Johnson, O. R. (1968). *Religion and our Schools*. London: Hodder & Stoughton.

Piaget, J. (1929). *The Child's Conception of the World*. Tomlinson, J. & A. (Trans.). London: Routledge & Kegan Paul.

West Riding Agreed Syllabus (1966). *Suggestions for Religious Education*. Obtainable from Education Dept., County Hall, Wakefield, Yorks.

Drugs Education

Questions and Answers

David A. Lane, Starcross School, Islington, London

Summary

A previous article (Lane, 1973b) demonstrated how two aspects of drugs education—help for individuals with their problems and education about drugs are often confounded. Having discussed the school's influence on individuals and their problems, this article focusses on education about drugs, and argues that there are many factors, such as the selective use by pupils of information, their attitudes and group membership, which must be taken into account if drugs education is to achieve anything.

How much awareness is there and how much education is taking place?

In 1969 it was estimated that almost half of the Medical Officers of Health in England and Wales denied knowledge of addiction or facilities to deal with the problem in their areas. The Advisory Centre for Education a year later (1970) produced a report showing a similar dismal state of affairs. Schools particularly seemed determined to deny the problem even when they knew of its existence.

Since 1969 there has been a rapid increase in awareness about drug problems but this awareness does not seem to have been matched by educational action. A content analysis of the reports of the Drug Dependency Discussion Group at the Hospital Centre (1968 to 1972) shows that most opinion is against the use of outside speakers to give occasional talks to schools on drugs, yet this same analysis shows that this is still the major educational method being used, and it also indicates that while many authorities are aware of the problem, most fail to take any comprehensive action to deal with it.

A recent report (1973) based on a national survey by the Hospital Centre, gives some grounds for optimism in that many more authorities, schools, colleges and universities are treading the path from ignorance to concern to action. But this same report gives evidence of the continuing inadequacy of educational provision since many people are still talking in terms of 'showing a film on the subject' as if this was the answer. Shamefully many colleges of education feel that the education content of their courses are a sufficient basis to the human relationships involved in such areas and a basis for counselling work. This is in spite of the fact that many seem unclear of the role of counselling even within their own organizations. The university departments of education seem less inclined to 'stand on their dignity' than the colleges and admit the inadequacies, so perhaps this more appropriate attitude might filter through. At the same time as authorities lack coordinated plans to deal with drug problems a fair amount of informal education in favour of drugs use is said to exist, and P.A.L. Chapple, Director of The National Addiction and Research Institute, maintains that this informal education has been very successful (1970). However, Wiener (1969) suggests that from an early age there is much anti-drugs education and since formal health education has little effect and most people have negative attitudes to drugs use, formal education should be avoided. Attitudes to drugs in general are confused. Thus, the attitude 'if you've got a pain take a pill' which leads to the development of the view of drugs as an answer to problems, and the existence of socially acceptable but dangerous drugs such as alcohol and tobacco alongside the belief that drugs are dangerous because a law says so, inevitably leads to educational difficulties. This confusion among specialists is evident in the attitudes of education authorities who believe alternatively that education

---

Acknowledgement

The author would like to thank Dr Anthony Wood of The Health Education Council for his advice on the preparation of this article.

about drugs is useful or does more harm than good (Report from the Hospital Centre, 1973).

This article attempts to answer this question.

The experimental study of drugs education, what does it tell us?

A study by the National Association of Youth Clubs indicates that the information given in drugs education lessons incites many people to want to try drugs who might otherwise not have done so. This problem of the arbitrary results of education and the difficulty in predicting its effects, is supported by other work in the mass communications field although the view that individuals are incited is more contentious.

McQuail (1969) has argued that there is overwhelming evidence that the net changes in attitudes or opinions as a result of persuasive material are likely to be small. The more likely effect of direct presentations of subjects such as drugs in which individual's attitudes, prejudices, and values are involved is their filtration through various other influences. Thus friends, opinion leaders and so on are more likely to influence the interpretation of the material received. Berelson and Steiner (1964) argue that the effects of communication tend to take the form of a reinforcement of existing attitudes and opinions, individuals hearing and seeing that which is favourable to their pre-dispositions. This is particularly important in the drugs field not only in terms of the potential user but also in terms of prejudice towards the user.

This filtration effect lies not only in individual responses to communications but also in the fact that the content of communications are often mediated by the structure of relations amongst the listeners, information flowing from source to opinion leaders and then to less active members of groups. Such evidence as there is suggests that communications are used by individuals to suit their own ends. The communicator is in a largely powerless position. This is particularly important in the area of drug use, since as has been argued in various studies, attitudes and values tend to be developed in intimate personal relationships (Sutherland and Cressey, 1955). Attitudes favourable to drugs use are likely to be learnt in cultures which favour their use (Becker, 1963).

It is not the communication which matters, but what recipients do with the communications they receive. In particular does the individual view himself as someone to whom a message is addressed? Such research as there is does not show that drugs education is or is not harmful but that it is likely to be used selectively by individuals. Thus the drug user may accept pro-drug education while the non-user may accept anti-drug messages. The person who does not care may ignore the message altogether, regarding it as something not addressed to him in spite of its preventive or informative aspects. We must therefore look at the groups to which communications are directed.

The dissemination of information

The particular structure of schools and colleges is such that they encourage deviant communication networks and the development of latent cultures (Lane, 1973b).

The school, like any organization, is constructed to achieve particular objectives but the achievement of these objectives is influenced by both formal and informal factors. Two factors which are important in this context are the structures of power and success within the school. Power is defined as the ability to affect the actions of others even against their will, and success as achievement in those activities designated as being of high status. Both of these to some extent are organizational demands which are external to the needs of the individual. However, these demands are not necessarily irksome because many pupils recognize and accept them as legitimate. There is congruence between the pupil and the school.

Pupils who find that they are unable to achieve the rewards offered or those for whom the rewards are meaningless may become dissatisfied and, feeling powerless to change the system, may adopt alternative deviant means to achieve their personal objectives either in isolation or as part of deviant, transitory, violent or drug-based groupings. This is just the sort of response that, in wider groupings, is outlined in the work of Merton (1949), Cloward (1959), Yablonsky (1962) and Cohen (1955). Because the school is organised in an age/ability/background system the formation of such groups is made easier. These groups are likely to evolve around hidden (latent) cultural influences. Drugs provide one such influence, which, because of the self-contained, if in part illusionary, rewards they offer, can become powerfully motivating. The influences satisfy many of the pupils' needs including a sense of achievement and belonging and provide a level of success unavailable in the school.

However, in addition to deviant behaviour patterns, deviant communication networks and language also develop and these influence the information to which the individual is exposed and consequently the effects of any education are limited by this network.

That language is based on social relationships is clear from the work of Halliday (1971), Goffman (1964) and several other researchers. Social groups reduce the external pressure on the individual, and because communication is a purposeful act, that form of communication best suited to needs is used, e.g., communication for the purpose of reinforcement of group and personal needs is based on conformity to the language of the group (Matza and Sykes, 1957; Bernstein, 1972). Labov (1969) has shown how the situation of communication between individuals from different groups is restricted.

The educator and the pupil are talking different languages. This is not simply a matter of codes but of the meanings attached to the language and the non-linguistic aspects of the situation. In a study of response to educational material in this subject area the author (1973c) has found that there are clear problems of language use involving difficulties such as inability to read the necessary material and also difficulties in understanding the points made in such lessons. This reaches the point where not only does the individual not co-operate with agencies external to himself but he does not see the need for such co-operation. The need for teachers to strive to maintain communication with such groups is clear (Lane, 1973a).

In any school class there are likely to be pro- and anti-drugs pupils whose minds are made up and who are unlikely to be influenced by education. These must be approached informally. For other pupils who are ambivalent or not committed, reasonable communication is likely to be most effective particularly when it shows both sides of the argument and does not involve high level threats of the dire consequences of drugs use. It is such two-sided communications which enable individuals to deal with contradictory information at a later date. See for example, a study of health education and of propaganda by Lumsdaine and Janis (1953) and Janis and Feshback (1953).

Unfortunately, many educational attempts in this area use exactly the opposite approach. The film 'Better Dead' is an example of this, high-threat, one-sided approach. The Institute for the Study of Drug Dependence (ISDD) has established a review panel for films, and teachers wishing to use films can obtain from the Institute copies of the comments of the panel on individual films. This same body in a recent survey (Dorn, 1972), found that pupils often knew more about the subject than their teachers. Teachers should therefore refer to reliable sources of information, such as the Media Resources Centre of the Inner London Education Authority, but even their film list should be checked against the ISDD review.

It is pointless convincing pupils that cannabis and heroin are both dangerous if when they leave the classroom they are faced with views they can't ignore which show cannabis not to be harmful. It is not important in communication terms whether or not this pupil-based information is correct only that it be believed. The danger is that having discovered that the teacher was 'lying' about cannabis the pupil may feel that perhaps other drugs also are not as harmful as the teacher made out. The status of the teacher, or those in authority generally, as a reliable source of information might be reduced very seriously by such contradictory communications.

The effects of such situations in the long run are difficult to assess but preliminary results from a study (1973c) undertaken by the author show that these differences do have long-term consequences for communication between teacher and pupil. And the ISDD study (Dorn, 1972) is being followed up on a large sample so data are beginning to be made available.

Interim summary

It has been argued then that communications are used selectively by individuals and interpreted in terms of their existing attitudes and group membership. Given this factor the aims of drugs education can now be considered.

The aims of drugs education

The aim of drugs education is not to prevent drug use, as this is impossible, but it is to enable individuals to make decisions based on realistic interpretations of evidence. The aims therefore are the same as in many other areas of education and do not require special justification. Drug use is part of our society and always has been therefore if one of the aims of education is to prepare individuals to take their place in society, education about drugs naturally forms part of this. To ignore it because it is unpleasant constitutes a neglect. Certainly, there are particular problems associated with drugs education, but the fact that it might be difficult is not an argument for not teaching it.

The most important of the problems on drugs education is created by the attempt to treat education about drugs as an exercise in preventive medicine. You cannot deal with the problem of dependency by talking about drugs. The two are separate.

Education about drugs concerns drug use as part of patterns of behaviour within society; it properly belongs in social science lessons where it can be viewed in perspective against a background of sociological theory on deviance, delinquency, etc.

Drugs education can also properly be taught as part of general health education in which the perspective of a whole range of health matters can be brought to bear. It is in this respect that the DES pamphlet 'Drugs and the Schools' (1972) makes its contribution, and although this pamphlet has been much criticized for its omissions it is nevertheless prime reading for teachers, particularly the final section which makes it clear that punitive action is unlikely to be helpful. Other aspects of drug use, such as the legal question, rightly form part of such an approach, but not as a sledgehammer to force a teacher's opinion on pupils. Education is not indoctrination.

Drug education has formed part of GCE syllabuses in sociology for some time, and taught as part of sociology it is effectively placed in a proper context. But if placed in a course on social problems as part of the non-social-scientific 'social studies' lessons taught in some schools then the fears of those who feel that it might do more harm than good are realized.

Treating drugs as part of social problems while seemingly reasonable may in fact create awareness of social pathology rather than an understanding of society. Pupils may be given the impression that society is falling apart under the pressure of divorce, drugs and delinquency. The emphasis in such courses is on deviance as if this were something apart from society as a whole, and it leads to the development of just those attitudes it is designed to prevent. A fourteen-year-old girl once said to me, 'Adults say we're all concerned about drugs and sex but here it's the teachers; they talk about nothing else, we get it in English, social studies, P.E., and even religion.' Because many people confuse giving information with helping individuals to learn to cope with the stresses of living they misguidedly slant the information they impart. Thus the content of education prevents the development of realistic attitudes about society. One example of this is the view of society's institutions as democratic and fair. When teaching concentrates on the way institutions are supposed to function rather than how they do in fact function, individuals are left with unrealistic ideas which leave them embittered when faced with the reality of a system which operates 'differential enforcement' and clamps down on those whose dress, opinions, etc., are not those of middle class respectability (Becker, 1963; Lambert, 1969).

By fostering an idealistic conception of the democratic society rather than preparing individuals to cope with it as it is, schools may encourage the problems which lead to deviance and confirm the opinion that adults never tell the truth except when it suits them.

Conclusion

At the beginning of this article it was promised that an answer to the question, 'Does drugs education do more harm than good?', would be attempted. In so far as this is possible on the basis of present evidence the answer would appear to be that present educational practice can do more harm than good because it enables individuals to confirm their existing prejudices and attitudes and does not prepare individuals to cope with stress. This failure is in part due to confusing teaching about drugs and the prevention of drug use.

However, if drugs education is treated reasonably as part of an overall educational policy it might still be useless as at this stage it cannot be proved that it is beneficial. Nevertheless, it may enable individuals to make realistic decisions and this can only be considered a benefit. Perhaps the greatest advantage in this approach is that the deviant may stop regarding his teachers as fools or liars.

Providing teachers think carefully about the aims and implications of their actions, and keep in mind the points made in this article, they can successfully undertake drugs education as part of education of the whole man, based firmly on fact and reason.

References

ADVISORY CENTRE FOR EDUCATION (1970). *Where on Drugs*. Cambridge.

Becker, H. S. (1963). *Outsiders: Studies in the Sociology of Deviance*. New York: The Free Press.

Berelson, B. and Steiner, G. (1964). *Human Behaviour: An Inventory of Scientific Findings*. New York: Harcourt Brace & World.

Bernstein, B. (1972). 'Social class, language and socialization', in Giglioli, P. P., *Language and Social Context*. Harmondsworth: Penguin.

Chapple, P. A. L. (1970). Reprint 472, Report of the drug dependency discussion group. London: The Hospital Centre.

Cloward, R. A. (1959). 'Illegitimate means, anomie and deviant behaviour', *Amer. Soc. Rev.*, **24**, 2.

Cohen, A. K. (1955). *Delinquent Boys: The Culture of the Gang*. New York: The Free Press.

Department of Education and Science (1972). 'Drugs and the Schools', available from the Department of Education and Science, 39 York Road, London, SE1.

Dorn, N. (1972). 'The ISDD drug education survey', *Drugs and Society*, **9**, 1, 21.

Goffman, E. (1964). 'The neglected situation', in Giglioli, P. P., *Language and Social Context*. Harmondsworth: Penguin.

Halliday, M. A. K. (1971). 'Language in the social perspective', *Educ. Rev.*, 23, 3.

Janis, I. L. and Feshback, S. (1953). 'Effects of fear arousing communications', *J. Abnorm. Soc. Psychol.*, 48, 78.

Katz, E. (1957). 'The two-step flow of communications', *Publ. Opin. Quart.*, 17, 311.

Labov, W. (1969). 'The logic of non-standard English', in Giglioli, P. P., *Language and Social Context*. Harmondsworth: Penguin.

Lambert, S. (1969). 'The police can choose', in *Crime, Deviance and Social Sickness*, a New Society Reader. London: New Society Publications.

Lane, D. A. (1973a). 'Report of a drugs conference', *Educational Bulletin*, 15, Bedfordshire.

Lane, D. A. (1973b). 'Individuals and systems: aspects of educational issues in drug dependence', *Educ. Res.*, **16**, 1, 52–7.

Lane, D. A. (1973c). 'Pathologies of communication', *Comm. Health*, 5, 3.

Lumsdaine, A. A. and Janis, I. L. (1953). 'Resistance to counter propaganda produced by one-sided and two-sided propaganda presentations', *Publ. Opin. Quart.*, 17, 311.

Matza, D. and Sykes, G. M. (1957). 'Techniques of neutralization', *Amer. Soc. Rev.*, **22**, 6.

McQuail, D. (1969). *Towards a Sociology of Mass Communications*. London: Collier Macmillian.

Merton, R. K. (1949). *Social Theory and Social Structure*. New York: The Free Press.

Reports of the Drug Dependency Discussion Group available from The Hospital Centre.

Report of the Education Working Party of the Drug Dependency Discussion Group available from The Hospital Centre.

Sutherland, E. H. and Cressey, D. R. (1955). *Principles of Criminology*. Philadelphia: J. B. Lippincott & Co.

Weiner, R. S. P. (1969). *Drugs and School Children*. London: Longmans.

Yablonsky, L. (1962). *The Violent Gang*. Harmondsworth: Penguin.

Film

'Better Dead', available from Guild Sound and Vision, Kingston Road, London, SW19.

Useful addresses

Teachers Council on Drug Education, 437, Royal Exchange, Manchester.

The Association for the Prevention of Addiction, 33 Long Acre, London, WC2.

The Institute for the Study of Drug Dependence. Now at: 3 Blackburn Road, London, NW6.

The Drug Dependency Discussion Group, The Hospital Centre, 24 Nutford Place, London, W1.

In Search of an 'Early Teaching Grammar'

John L. Presland, Educational Psychologist, Warwickshire County Council

Summary

Research on the psychology of learning suggests that identifying, labelling, and constructing 'types' of sentence could help improve children's punctuation and sentence construction. Samples of written work from a primary school were, therefore, analysed into 'sentences', and these classified into a manageable number of structural patterns which could be given labels meaningful to young children. The most common pattern was analysed further to yield sub-patterns which could be arranged into an appropriate teaching sequence. The results showed that an 'Early Teaching Grammar' was feasible. The implications of this for practical teaching and for further research, are discussed.

Introduction

The investigation described here was carried out in the hope that some kind of classification of children's early written English could be found which would help them to learn the following skills more efficiently.

1. *Punctuation*

The earliest written work is simply speech transferred to paper. If it is to make sense, capital letters and full-stops must be added to divide it into sentences. Odom (1961, 1962, 1964) found, with American children, that these simple punctuation skills were not developed, even to the criterion of 'right half the time', until about the age of ten, and there seems no reason to suppose that children in this country are much more advanced. Do we need to wait so long?

In order to punctuate sentences, the child must be able to *identify* them, and this is very difficult. All definitions of a sentence known to the author are either invalid or useless. Such definitions as 'a group of words which makes sense' and 'a complete thought' would apply just as much to 'A man walking' written as a label for a picture as to 'A man is walking'. The reason why the latter is normally called a sentence, and the former not, lies in their grammatical structures. There are many types of grammatical structure which constitute a sentence and it is most unlikely that they can all be subsumed under any one definition useful to young children. Nor is it likely that they can be identified by pauses in speech, as is sometimes suggested; Wilkinson (1971) gives a useful account of the relevant differences between written and spoken language. It may, however, be possible to teach children to recognize the particular types, and identify them by labels which are meaningful to them. They could then detect examples in their written work, and punctuate them accordingly. Later, perhaps, they could identify them as they are composed, and punctuate as they write.

2. *Improving Sentence Structure*

Early investigations, summarized by McCarthy (1954), and subsequent work by Hunt (1965), O'Donnell *et al.* (1967), and Zeman (1969), have shown that children's earlier written language structures are simpler than those which come later. It has also been found that gifted children of about eight can be taught to increase the complexity of sentences in their written work (Hilfman, 1970). Miller and Ney (1967) found the same with average ten-year-olds: an experimental group used more of the structures taught in their written work than did a control group. Possibly younger children could begin by constructing very simple types of sentence and be helped to alter and expand them in a variety of ways.

Research on the psychology of learning suggests several reasons why the identification, labelling, and constructing of definite 'types' of sentence should help in teaching the skills above:

1. *It helps certain aspects of the learning to be verbalized*

Naming elements in the learning process, and putting learning strategies into words, can both increase learning efficiency. Evidence supporting these assertions can be found in studies by Luria and Yudovich (1959), Luria (1961), Frandsen (1967), Caldwell and Hall (1970), Davids and Adams (1970), MacGinitie (1969), Otto *et al.* (1971), Lewis (1963), and Stones (1970). It follows that giving a verbal label to a sentence pattern should help its identification, and form a basis for verbalizing the process by which the pattern is altered and expanded.

2. *It allows a logical arrangement of material which helps to make the learning process meaningful*

That relating the elements of a learning task in a meaningful way aids children's learning has been demonstrated by Brownell and Moser (1949) and by Fuld (1970). It has been shown with such disparate groups as first-year grammar school boys (Skemp, 1962) and subnormal boys (Jones *et al.*, 1972) that expanding such relationships into a logical system is also effective. Menyuk (1969) found that such a system underlies pre-school *oral* language, though no doubt it is used unconsciously. It may well be, therefore, that the conscious application of a logical system to early written language can be helpful.

3. *It facilitates transfer of learning*

Transfer is said to have occurred when learning one task makes the learning of a subsequent different task easier. This is more likely to occur if basic principles can be mastered in the first task and then applied to the second (Samuels and Jeffrey, 1966; Klausmeier, 1969; McCutcheon and McDowell, 1969; Caldwell and Hall, 1969). In the case of early written English, identifying, labelling, and constructing sentences of a certain type allows formulation of principles which can then be applied to a vast range of other sentences of the same type.

4. *It helps the teacher to guide the child's learning*

Evidence summarized by Frandsen (1967), Ausubel and Robinson (1969), and Leith (1965) indicates that guidance is important for learning as well as the currently popular discovery procedures. It seems likely, indeed, that young children may need a great deal of guidance before they can make appropriate discoveries. Some evidence (Mednick and Lehtinen, 1957; Goulet, 1968) indicates that they generalize their learning more than older children and are therefore likely to use what they have learned in the wrong as well a the right circumstances, unless guided. Other evidence (Goulet, 1971) suggests that, without such guidance, they are more likely to be distracted by potentially misleading elements in a learning situation. It has already been pointed out that a classification of sentences into 'types' helps the learning to be verbalized and enables a logical approach to the learning. These are both factors which help the teacher to provide guidance. The potential of such a scheme for facilitating transfer of learning can also contribute to guidance.

Research into the use of grammar in English teaching seems confined to pupils of secondary school age, and even there the results are discouraging. The position has been well summarized by Braddock (1969). Teaching of traditional grammar, the kind which most of us are familiar with from our schooldays, has been found to have no effect on punctuation or sentence structure (Hunnicutt and Iverson, 1958). More recently developed grammars have also been investigated to a small extent. Mixed results have been obtained with the Structural Grammar largely originated by Fries (1952), and the evidence for Transformational Grammar, of which Chomsky (1957) is the main architect, is mildly encouraging. These researches have, however, been carried out with near-adults, and the grammars are so complex that it is hard to see how they could be used with young children. Perhaps more hopeful is the approach described by Fraser (1969), who argues that previous teaching of grammar has concentrated too much on analysis of sentences and too little on their production. He describes the use of traditional grammar to *construct* sentences and claims it has been successful with slower secondary modern and with junior children. No controlled experiments are, however, reported, and it is likely that even this approach would be too difficult for younger children. It seems that we need to start again, and to look at young children's writing and try to identify grammatical patterns which are sufficiently simple and meaningful to be taught to and used by such children. What follows is an attempt to identify them, to estimate their relative frequencies, and to analyse them in various ways which might suggest a sequence for teaching.

Method

The head teachers of a variety of schools were sent a circular asking for samples of written work marked with the age and sex of the pupils. It was stressed that the object was to obtain knowledge about the types of sentences children were interested in writing rather than about their skill in writing them. It was, therefore, quite acceptable for a child to have copied his sentences from the teacher's writing, so long as it was the child who had thought up the sentence to be written. The type of written work was not specified in any way. Samples were received from a number of schools, and the contributions from one of them were chosen for a first analysis, purely on the ground that it was the only school producing samples from every child present over a wide age range. It was a Roman Catholic junior and infant school drawing its pupils almost exclusively from municipal housing. The head teacher felt that the general intellectual and educational level of the pupils was below the national average. Bright children were uncommon. The samples were on a wide range of topics, but the recurrence of certain subjects suggested that some teachers had offered suggestions to choose from.

Since the samples were to be grouped by age for analysis, and there were very few from eleven-year-olds, these were discarded. The remaining samples, from children of six to ten years, were then divided into units which could be regarded as the sentences children would write (or combine into compound sentences) if they could punctuate correctly. Each unit consisted of a main clause, and any subordinate clauses depending on it. This method was adopted because Loban (1963), who called it the 'communication unit', and Hunt (1965), who called it the 'T-unit', found it useful in their studies of children's language. It will here be referred to as the 'T-unit'. The following sample shows how the division was made:

I like branded because they always win / and it is fun / and I like it because sometimes he gets killed.

Collections of words which could not be fitted into this analysis were excluded from consideration. This was sometimes because they did not make sense and sometimes because they were records of speech. Both occurred very infrequently. Another problem arose from the word 'and', which occurred excessively. Hence many T-units began with this word and would strictly, therefore, be incorrect sentences. It was felt that this should not be allowed to interfere with the analysis, but simply be remembered when any results were applied to teaching.

The divided samples were analysed in the following ways:

1. They were examined to identify sentence patterns which could be given a label meaningful to young children. The number of times each pattern recurred at each age level was determined, and converted into a percentage of the total number of T-units at that age.

2. The most common pattern identified in (1) was further analysed into sub-patterns, and their frequency was recorded as a percentage of the total number of T-units conforming to that pattern at each age.

3. The most common sub-pattern identified in (2) was analysed to obtain the percentage distribution of different subjects of T-units.

4. All T-units conforming to the most common pattern identified in (1) were examined to determine what percentage began directly with their subjects, and what percentage began in each of several other ways.

5. A further analysis was made of the most common pattern to determine the percentages of T-units conforming to it, which:

- (a) had conjunction of main clauses (*I got out my ray gun and shot him; I had my tea, went out and played in the park*);
- (b) had subordinate clauses (**When I was at Blackpool,** *I went on a lorry*);
- (c) had adjectival or adverbial phrases not embedded in subordinate clauses (**One day** *Ned fell asleep* **in the sun**);
- (d) had none of the foregoing complications, had a direct and/or indirect object which was not a subordinate clause, and did not begin with the subject (*and we had a nice time*);
- (e) were the same as (d), except that they began with the subject (*I went on the ladder*);
- (f) had none of the previous complications except that they did not begin with the subject (*and they go fast*);
- (g) were the same as (f), except that they began with the subject (*The trains go very fast*).

6. All verbs occurring more than twice in T-units of the most common pattern at ages six

and seven were recorded. The number of times each occurred at each age was noted.

Results

The number of samples at each age, and the number of T-units into which they were divided, were as follows:

Age 6	31 samples	125 T-units
Age 7	34 samples	118 T-units
Age 8	45 samples	327 T-units
Age 9	39 samples	297 T-units
Age 10	44 samples	315 T-units

The results of the various kinds of analysis described above are recorded below.

1. The following sub-patterns were considered suitable. Though each is put into a particular number and tense, other numbers and tenses, and also negatives, are included, e.g., the *somebody did something* sentence includes *somebody does something*, *some people do something*, and *somebody did not do something*. Sentences with non-human subjects are also included.

i. The 'somebody was something' sentence, e.g., *I was very scared; I am a girl.*

ii. The 'somebody did something' sentence, e.g., *I ran to the cellar*. Included in this category were T-units with the same grammatical structure where it was arguable whether anything was actually being 'done', e.g. *I like branded; I had one hour to find the bomb.*

iii. The 'somebody was doing something' sentence, e.g., *I was watching golf.*

iv. The 'somebody likes doing something' sentence.

v. The 'somebody has done something' sentence.

vi. The 'somebody will (or shall) do something' sentence.

vii. The 'somebody had to do something' sentence.

viii. The 'somebody can do something' sentence.

ix. The 'there was something' sentence, e.g., *There was once a very old house.*

There were also a few other patterns, but none of them occurred more than three times in the whole sample.

Table 1 shows the percentage occurrence of each pattern to the nearest whole number. It was notable that the *somebody did something* pattern accounted for over half the total number of T-units at all age levels; that *somebody did something* and *somebody was something* together accounted for 78 per cent or more of the total T-units at all levels; and that 95 per cent or more of the T-units at each age were accounted for by the nine sentence patterns identified.

2. The frequency at each age of each sub-pattern of pattern (*ii*), expressed as a percentage of the total number of T-units conforming to that pattern is, shown in Table 2. 'Unclassified' refers mainly to sentences generally recognizable as pattern (*ii*), but not entirely comprehensible. It was notable that over half the total number of examples of the *somebody did something* pattern at each age were accounted for by the *somebody did something* sub-pattern; that this sub-pattern, *somebody had something*, and *somebody does*

Table 1: Occurrence of sentence patterns as a percentage of total T-units at each age

Sentence pattern	*Age 6*	*Age 7*	*Age 8*	*Age 9*	*Age 10*
Somebody was something	14	8	19	18	13
Somebody did something	71	81	59	61	70
Somebody was doing something	7	5	6	5	4
Somebody likes doing something	5	1	0	–	–
Somebody has done something	1	1	2	2	1
Somebody will (or shall) do something	1	–	2	2	–
Somebody had to do something	1	–	1	2	3
Somebody can do something	1	2	2	3	2
There was something	–	2	3	5	1
Others	–	1	5	2	2

Table 2: Occurrence of sub-patterns as a percentage of total number of 'somebody did something' sentence patterns at each age

Sub-pattern	*Age 6*	*Age 7*	*Age 8*	*Age 9*	*Age 10*
Somebody did something	75	85	63	86	81
Somebody had something	14	7	6	5	10
Somebody does something	9	2	18	2	5
Something did something	–	3	4	3	2
Somebody has something	2	2	2	1	1
Somebody does not do something	1	–	–	1	–
Something has something	–	–	2	1	–
Somebody did not do something	–	–	1	1	–
Something does something	–	–	4	–	–
Something had something	–	–	1	–	–
Somebody had done something	–	–	–	1	0
Unclassified	–	–	1	1	–

NOTE. 'Somebody' includes animals, since these tend to be personified by young children.

something together accounted for over 85 per cent at each level; and that 99 or 100 per cent of the examples of the pattern at each age level were accounted for by the eleven sub-patterns identified.

3. Table 3 shows the percentage of pattern (*ii*) T-units with each subject at each age. Most T-units in the 'others' category had more than one subject (*One day me and my friend got into trouble*). *I* was by far the most common subject at all ages.

4. The percentage of pattern (*ii*) T-units beginning in each of a number of ways at each age appears in Table 4. 'Others' consisted largely of adverbs, adverbial phrases, and adverbial clauses, mostly indicative of the time at which something was done. It was clear that the subject and the word *and* were the most common beginnings at ages six to nine. At age ten the 'others' category also became very common.

5. Table 5 shows the percentages of pattern (*ii*) T-units of various levels of complexity at each age. It was noteworthy that conjunction of main clauses was rare throughout; that subordinate clauses were rare at the earlier ages, but more common later; that adjectival and adverbial phrases and objects occurred commonly at all ages; that there were few T-units at any age without one or more of the above complications; and that of the T-units without conjunction of main clauses, subordinate clauses, or adjectival or adverbial phrases, the majority did not begin with the subject (except at age eight).

The most common verbs in pattern (*ii*) T-units at ages six and seven, and their frequencies, were

Age 6: *went* (29), *had* (13), *saw* (10), *brought* (4), *watched* (3), *like* (3), *came* (3).

Age 7: *went* (34), *had* (9), *played* (8), *bought* (6), *saw* (4), *got* (4), *brought* (3), *came* (3).

Conclusions

This investigation was designed, firstly, to test the feasibility of an early teaching grammar, and the results suggest that such a development is a genuine possibility. In the sample of written work studied, 95 per cent or more of the T-units could be accounted for by only nine sentence patterns. Perhaps, if children could be taught to identify these patterns, they could subsequently learn to locate and punctuate them in written material, and to use them as a basis for improving the structure of the sentences they write. The relative frequencies of the different patterns, and of different sub-patterns and elements within each pattern, could suggest an order of teaching, the most common being taught first. This is discussed in more detail in the 'suggestions for teaching' section below.

Table 3: Percentage of 'somebody did something' sentence patterns with each subject at each age

Subject	*Age 6*	*Age 7*	*Age 8*	*Age 9*	*Age 10*
I	59	74	43	37	41
We	9	5	17	8	23
He	2	2	14	12	7
She	3	4	8	4	4
They	–	–	2	8	7
Proper noun	9	1	7	3	5
Article and noun	12	1	7	15	4
My something (mostly a relative)	3	10	2	4	3
Others	3	4	–	8	5

At this stage, however, conclusions can be tentative only. A relatively small sample of written work was studied and it was from only one school. A larger sample from a variety of schools could give different results. Furthermore, there is no direct evidence to show that teaching techniques based on the findings would actually work. Research is needed in both of these areas before the results obtained here can be applied confidently in the classroom. If such research confirms the present results, further analysis of written work of a wide variety of types would be justified, so that a detailed grammatical system could be devised, and further teaching techniques worked out and tested. Perhaps, eventually, a system would emerge which was a reasonable and comprehensive description of the structure of children's early writing. It should then be possible to establish norms for the development of grammatical structures to help in assessment and further research, though such norms would themselves be susceptible to modification by changes in teaching.

Suggestions for teaching

The first stage in applying the present findings in the classroom might be to teach children to identify the sentence patterns, at first orally, and then in written form. They could then learn to locate and punctuate them in written material. At first this could be material specially written by the teacher to include sufficient examples of the sentence pattern being taught; then writing the children had produced in response to instructions designed

Table 4: Percentage of 'somebody did something' sentence patterns with each type of beginning at each age

Beginning	*Age 6*	*Age 7*	*Age 8*	*Age 9*	*Age 10*
Subject of T-unit	40	32	46	34	29
And	42	56	28	37	22
Then	7	5	3	3	12
But	2	–	1	2	2
So	–	1	6	5	6
Others	9	6	16	19	29

Table 5: Percentage of 'somebody did something' sentence patterns of various levels of complexity at each age

Level of complexity	*Age 6*	*Age 7*	*Age 8*	*Age 9*	*Age 10*
Conjunction of main clauses	1	4	4	7	4
Subordinate clauses	8	2	12	20	19
Adjectival or adverbial phrases	34	24	29	29	41
None of above, but with object (not beginning with subject)	35	40	20	21	25
None of above, but with object (beginning with subject)	17	23	28	13	11
None of above (not beginning with subject)	2	5	6	11	7
None of above (beginning with subject)	4	1	3	3	1

to elicit the required patterns [e.g., 'Write sentences to tell all the things you did last night' for eliciting pattern (*ii*)] ;and, finally, ordinary written work.

To ensure that the teaching were maximally useful to the children, the most commonly used sentence patterns would need to be taught first. In the written work studied, the *somebody did something* pattern was so much more common than any of the others at every age level that its confirmation by further research as the first to be taught seems very likely. *Somebody was something* may be the pattern to be taught next, and *somebody was doing something* could eventually be shown to be the next most common.

The *somebody did something* sentence can be used as an example of how the teaching of each pattern might be organized. It would be best to begin with the simplest possible examples likely to be useful for applying to written work. *I jumped* is one of the simplest one could imagine, and it may be best to start with sentences like it, though further developments would be necessary before application to written work. As Table 5 shows, the simplest category was very poorly represented in the written work studied, and most T-units even in that category were more complex than *I jumped*. It may, however, eventually be possible to narrow down considerably the types of sentences which can be looked for first in written work—in the present study, for instance, the great majority of pattern (*ii*) T-units were in the past tense, and referred to *somebody* rather than *something* (Table 2). *I* was by far the most common subject (Table 3). Table 4 shows that a sufficiently large proportion began with the subject to make it worthwhile excluding those that didn't until later stages of teaching, if the same result is found in further research. Table 5 suggests the possibility that conjunction of main clauses, subordinate clauses, and adjectival and adverbial phrases might be similarly disregarded at first. It was found also that *went* was by far the most common verb. The first sentences to punctuate in written work might, therefore, be those with no conjunction of main clauses, no subordinate clauses, and no adjectival or adverbial phrases, but beginning with the subject *I* and having *went* (i.e., the past tense of *go*) as the verb. Examples are: *I went out; I went to the seaside; I went on a slide; I went quickly to the big boy*. These could all be presented to the child, and he could be told that they are all *I went somewhere* sentences. He could then be presented with other sentences and set to practise deciding whether they do or do not conform to the same pattern.

Once this first pattern is mastered, other commonly occurring verbs indicating action might be introduced, such as *played*, *bought*, *brought*, *came*, and *watched*. The child could now be told that they are all *I did something* sentences. Then verbs not obviously indicating action (*had*, *saw*), could be brought in, the child perhaps being told that they are 'very like' *I did something*. Other subjects could later be used, and the full *somebody did something* generalization established. Non-human subjects could be left till last.

When identifying these early patterns in written work, problems would arise from the long strings of 'sentences' joined by *ands* which children so commonly write. Table 4 shows that a very large proportion of *somebody did something* T-units began with *and* at the earlier ages. One way of overcoming this might be to include sentences beginning with *and* as examples of the early patterns, and punctuate accordingly. Another might be to cross out all *ands* and begin each sentence with the word that follows. This would eliminate all compound sentences, which might be a useful corrective at this stage. Words like *then*, *but*, and *so*, more acceptable at the beginning of a sentence, could be dealt with by eventually introducing sentences beginning with them as examples of early patterns.

Further progress might be associated with a programme for improving sentence structure. Examples of simple sentence patterns could be expanded, at first in oral work, by asking such questions as, 'What sort of person did it?'; 'When did he do it?'; 'How did he do it?'; 'Why did he do it?'; This should result in inclusion of adjectives and adverbs, and clauses and phrases with the same functions. These more complex patterns could subsequently be identified and punctuated in written work. Conjunction of main clauses would need to come eventually, and sentences to be joined together to form compound sentences.

The teaching techniques suggested would need, of course, to be related to the teaching of written English as a whole. Writing is a complex activity, in which a child makes use of a variety of skills. He must formulate worthwhile ideas to write about, put them into meaningful language, write letter-forms legibly, spell words correctly (or at least recognizably), and ideally, produce grammatically correct, properly punctuated sentences organized into an effective whole. Learning any one of these skills is simplified if, initially, it is taught in isolation from the others, but the long-term aim must be to integrate them into the total writing process. The most inspired ideas, for instance, will not help the child to write if he lacks the skills needed to transfer them to paper; nor is there much point in possessing the latter skills if the child has no ideas to which he can apply them. It is suggested, therefore, that teaching should begin by helping the child generate ideas and formulate them in spoken language. Some of his spoken language is eventually written down by the teacher. Later the child attempts to copy some of the writing produced by this dictation, which provides an opportunity to begin the teaching of handwriting. Spelling soon follows, and the child is gradually encouraged to embody some of his ideas in a little writing of his own, the demands for accuracy in writing skills being minimal. More advanced handwriting and spelling skills are later taught systematically, and gradually introduced into the child's writing. When, eventually, punctuation and grammatical accuracy are dealt with in the same general way, the teaching techniques proposed earlier could be relevant.

Implicit in the above discussion is the problem of introducing so many different skills into written work without destroying its fluency or allowing concentration on one skill to interfere with efficiency in another. If, for instance, great emphasis is placed on punctuation and grammatical accuracy, it seems likely to inhibit the flow of the child's ideas; while concentrating too heavily on development of ideas may reduce the accuracy of punctuation and grammar to a level at which his written record does not communicate the ideas meaningfully. Some kind of compromise seems needed, and there are wide differences of opinion, which existing research does not really help to resolve, as to where the balance between different skills should lie. One useful compromise could be to concentrate on introducing one skill only in a particular piece of writing. If we are concentrating on formulation of ideas, we do not expect too much accuracy in punctuation and grammar; if we are concentrating on the latter, we do not expect the ideas to be quite so inspired. When applying the skills learned to writing for some purpose other than learning how to write, the distinction made by Burrows *et al.* (1964) between practical and personal writing seems helpful. The former is intended as a means of communicating with others, and therefore, requires accuracy in most aspects of written English, including punctuation and grammar. Personal writing, on the other hand, is written for the child's own satisfaction, and as his personal record, which need not be properly punctuated or grammatically correct. Even in the case of practical writing, however, it may be best for the child to produce a first draft in which he concentrates mainly on formulating his ideas, and to correct this with his teacher's help, and perhaps some time after its production, before writing an accurate final version. A familiarity with the sentence patterns identified in this article could then be useful in improving punctuation and sentence structure during the correction process.

It must be reiterated that the proposals here have been limited to those deriving from the analysis of written work described, which was limited in scope. Much more research is needed before their applicability and effectiveness in the classroom can be confirmed, and their relationships with other skills and teaching techniques established.

References

AUSUBEL, D. P. and ROBINSON, F. G. (1969). *School Learning: An Introduction to Educational Psychology*. New York: Holt, Rinehart & Winston.

BRADDOCK, R. (1969). 'English composition', in EBEL, L. (Ed.) *Encyclopaedia of Educational Research* (4th ed.), 443–61. London: Macmillan.

BROWNELL, W. A. and MOSER, H. E. (1949). *Meaningful Versus Mechanical Learning: A Study in Grade Three Subtraction*. Research Studies in Education 8. Durham, N. Carolina: Duke University.

BURROWS, A. T., JACKSON, D. C. and SAUNDERS, D. O. (1964). *They All Want to Write: Written English in the Elementary School* (3rd ed.). New York: Holt, Rinehart & Winston.

CALDWELL, E. C. and HALL, V. C. (1969). 'The influence of concept training on letter discrimination', *Child Devel.*, **40**, 1, 63–71.

CALDWELL, E. and HALL, V. (1970). 'Concept learning in discrimination tasks', *Devel. Psychol.*, **2**, 1, 41–8.

CHOMSKY, N. (1957). *Syntactic Structures*. The Hague: Mouton & Co.

DAVIDS, R. E. and ADAMS, J. F. (1970). 'Verbal and imagery processes in children's paired-associate learning', *J. Exp. Child Psychol.*, **9**, 429–35.

FRANDSEN, A. N. (1967). *Educational Psychology* (2nd ed.). New York: McGraw Hill.

FRASER, H. (1969). 'The teaching of writing', in FRASER, H. and O'DONNELL, W. R. (Eds.) *Applied Linguistics and the Teaching of English*, 120–39. London: Longmans.

FRIES, C. C. (1952). *The Structure of English*. London: Longmans Green.

FULD, P. A. (1970). 'Syntactical mediation of paired-associate learning as a function of age', *J. Exp. Child Psychol.*, **10**, 2, 248–56.

GOULET, L. R. (1968). 'Verbal learning in children. Implications for developmental research', *Psychol. Bull.*, **69**, 5, 359–76.

GOULET, L. R. (1971). 'Optimum learning and optimal environments. A view from a child development perspective', *Ed. Tech.*, **11**, 2, 13–18.

HILFMAN, T. (1970). 'Can second-grade children write more complex sentences?', *El. Engl.*, **47**, 2, 209–14.

HUNNICUTT, C. W. and IVERSON, W. J. (1958). *Research in the Three R's*. New York: Harper.

HUNT, K. W. (1965). *Grammatical Structures Written at Three Grade Levels*. Research Report 3. Champaign, Illinois: NCTE.

JONES, E. C., MYERS, D. G. and CLEVELAND, D. L. (1972). 'Schema learning in retarded children', *Psychol. Reps.*, **30**, 2, 447–53.

KLAUSMEIER, H. J. (1969). 'Transfer of learning', in EBEL, L. (Ed.) *Encyclopaedia of Educational Research* (4th ed.), 1483–93. London: Macmillan.

LEITH, G. O. M. (1965). *Learning Theory in Relation to the School Situation*. University of Birmingham Department of Education.

LEWIS, M. M. (1963). *Language, Thought and Personality in Infancy and Childhood*. London: Harrap.

LOBAN, W. D. (1963). *The Language of Elementary School Children*. Research Report 1. Champaign, Illinois: N.C.T.E.

LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour*. Oxford: Pergamon.

LURIA, A. R. and YUDOVICH, F. I. (1959). *Speech and the Development of Mental Processes in the Child*. London: Staples.

MACGINITIE, W. M. (1969). 'Language development', in EBEL, L. (Ed.) *Encyclopaedia of Educational Research* (4th ed.), 686–99. London: Macmillan.

MCCARTHY, D. (1954). 'Language development in children', in CARMICHAEL, L. (Ed.) *Manual of Child Psychology* (2nd ed.), 492–630. New York: John Wiley.

MCCUTCHEON, B. R. and MCDOWELL, E. E. (1969). 'Intralist similarity and acquisition and generalization of word recognition', *Read. Teach.*, **23**, 2, 103–8.

MEDNICK, S. A. and LEHTINEN, L. (1957). 'Stimulus generalization as a function of age in children', *J. Exp. Psychol.*, **53**, 3, 180–3.

MENYUK, P. (1969). *Sentences Children Use*. Research Monograph 52. Cambridge, Mass.: M.I.T. Press.

MILLER, B. and NEY, J. (1967). 'Oral drills and writing improvement in the fourth grade', *J. Exp. Educ.*, **36**, 1, 93–9.

ODOM, R. (1961). 'Sequence and grade placement of capitalisation skills', *El. Engl.*, **38**, 2, 118–21.

ODOM, R. (1962). 'Sequence and grade placement of punctuation skills', *Calif. J. Educ. Res.*, **13**, 4, 179–85.

ODOM, R. (1964). 'Growth of a language skill: punctuation', *Calif. J. Educ. Res.*, **15**, 1, 12–17.

O'DONNELL, R. C., GRIFFIN, W. J. and NORRIS, R. C. (1967). *Syntax of Kindergarten and Elementary School Children: A Transformational Analysis*. Research Report 8. Champaign, Illinois: N.C.T.E.

OTTO, W., BARRETT, T. C., SMITH, R. J., DULIN, K. C. and JOHNSON, D. D. (1971). 'Summary and review of investigations relating to reading, July 1st, 1969 to June 30th, 1970', *J. Educ. Res.*, **64**, 6, 242–68.

SAMUELS, S. J. and JEFFREY, W. E. (1966). 'Discriminability of words and letter cues used in learning to read', *J. Educ. Psychol.*, **57**, 6, 337–40.

SKEMP, R. (1962). 'The need for a schematic learning theory', *Br. J. Educ. Psychol.*, **32**, 2, 133–42.

STONES, E. (1970). 'Verbal labelling and concept formation in primary school children', *Br. J. Educ. Psychol.*, **40**, 245–52.

WILKINSON, A. (1971). *The Foundations of Language*. London: Oxford University Press.

ZEMAN, S. (1969). 'Reading comprehension and writing of second and third graders', *Read. Teach.*, **23**, 2, 144–50.

Scoring Written Work in English as a Second Language

the Use of the T-unit

C. Gipps and E. Ewen, National Foundation for Educational Research

Summary
For the purposes of scoring essays written in a second language, two of the most important considerations are the intelligibility and the structural complexity of the writing.

Various disadvantages are inherent in the use of clauses and/or sentences as a basis for analysing structure in written work; a more satisfactory technique was developed by Kellogg W. Hunt in America, using what he termed a 'minimal terminable unit' or 'T-unit'.

This technique was applied in the scoring of the NFER open-ended writing and speaking tests, which formed part of the battery of 'Tests of English Proficiency for Immigrant Children'.

During the development of these tests, the battery was administered to Asian children, for whom it was found that the average T-unit length in writing and speech increased with increasing length of stay in Britain. (Average T-unit length has been found by Hunt and O'Donnell to increase with age for children writing and speaking in their native language.) The results of the NFER testing indicated that much of the development of proficiency in both the speech and writing of the Asian children tested took place after three-and-a-half years in Britain. (These findings were essentially a by-product of test development, and therefore must be viewed with caution.)

Scoring methods

In the construction of tests of writing ability, the relative value of open-ended versus objectively scored responses has been the subject of much controversy. The result has generally been that the written composition is still widely viewed as one of the most valid means of assessing writing ability, despite the inevitable complications which arise in the development of scoring procedures. An essay task was therefore included as part of the writing battery in the series of tests of English proficiency for immigrant children, recently developed by the NFER.

A written composition may be judged according to many different criteria, and the total score given to an essay is usually derived from a combination of measures of specific aspects of the writing, such as for example, length, accuracy, variety of content, linguistic maturity or literary style. In developing a suitable scoring system, the relative weight given to each of these aspects will depend to a large extent on the aim of the test, and on the individuals being tested. If the essay task is to be given to adults for example, to assess the development of writing skill in their native language, such factors as style and content will play a large part in the scoring of the essays.

In the NFER tests for immigrant children of junior school age the essay task is used to test writing ability in a second-language, and, as such, is concerned primarily with the capacity for effective use of the second-language to communicate in the medium of writing, rather than with the content or style of the communication itself. The effectiveness of written communication, in this sense, depends largely on its intelligibility, involving a minimal degree of grammatical accuracy. An effective communication could also be said to be one where optimum use is made of the structural features of a language (i.e. phrases, clauses and sentences), to maximize the information expressed in a given number of words, and so it was felt to be appropriate to include in the scoring system some measure of structural complexity.

It has been the standard practice in language development studies to assess structural complexity by calculating the mean sentence length and the mean clause length from which the subordination index, i.e. the ratio of dependent clauses to all clauses, is calculated. The mean sentence length, and the subordination index together provide a measure of structural complexity.

However, there are certain problems associated with this method:

1. The use of sentences is often unsuitable in the analysis of the writing of young children, where very little punctuation may be used, making the precise limits of a sentence difficult to define, especially in the case of primary school written work where punctuation is often under-emphasized in order that creativity and imagination may come to the fore.

2. The tendency which children have to join together a series of main clauses by the repeated use of the conjunction 'and' could have a most distorting effect, if sentence length were to be used as a measure of skill in writing.

3. There are several different definitions of what constitutes a clause.

The T-unit

In 1965 an American, Kellogg W. Hunt, carried out a study of samples of the writing of children who were native speakers of English.

In his study, Kellogg Hunt analysed 1000-word samples of free writing produced during the school year as a part of normal class work. The sample consisted of 9 boys and 9 girls of average intelligence (as measured by the California Test of Mental Maturity) from each of three age-groups; eight-year-olds, 12-year-olds and 16-year-olds.

He found that, although the average sentence length for each group showed an increase with age, there were nevertheless many individual differences, and he concluded that punctuated sentence length was only likely to be a satisfactory index of linguistic maturity for adults or for very able children and not for young children of average ability.

Hunt did find, however, that the mean clause length increased with age and that the average main clause had more subordinate clauses attached to it, i.e. the subordination index rose.

He therefore considered that a combination of these two factors—clause length and subordination—might prove a better measure of maturity, and was thus led to develop the idea of the T-unit or 'minimal terminable unit', as a means of dividing up continuous pieces of writing in such a way as to reveal the structural complexity.

A T-unit is defined as 'one complete main clause and any other clauses dependent upon it'. Such statements constitute the smallest unit in a communication that can be punctuated with terminal marks at both ends, hence the name 'minimal terminable unit'.

Hunt found that the mean T-unit length was more related to age than were mean sentence length, subordination index, or mean clause length: the longer the T-unit the more complex the writing, either because of increased clause length or increased subordination. Hunt found that with the younger children this was due to increased clause length, but with the older children it was due to increased subordination.

The use of the average T-unit length, as an index of maturity, has the advantage of accounting for all the subordination in the written work, and all the co-ordination between words, phrases and subordinate clauses, although it does not show the co-ordination between main clauses

The following are examples of T-units: (/ = end of T-unit)

1. At school we play football/
   (Main clause)
2. At school we play football, whenever we can/
   (Main clause) [Dependent clause (subordinate)]
3. At school we play football,/ but we never play netball.
   (Main clause) (Main clause)
4. At school we play football, and go swimming/
   (Main clause) [Dependent clause (co-ordinate)]

N.B. Each T-unit can contain one, and only one, complete main clause.

Since T-units, rather than sentences or clauses, had been used successfully in dividing up the written work of native speakers thereby providing a built-in measure of linguistic maturity, it was decided to experiment with these units in the scoring procedures for the NFER open-ended writing tests as a means of assessing the structural complexity of essays written in a second language.

Scoring system used in the NFER tests

The children were required to write an essay on one of the following three subjects: 'My Family', 'Coming to Britain' or 'The Fire'. For each essay, a cut-off point was set after approximately the first 100 words, at the end of the nearest complete T-unit. All the writing up to that point was then split up into T-units, any totally unintelligible tracts being crossed out. The average T-unit length was calculated by dividing the number of words by the number of T-units which they formed.

Having established a means of assessing the structural complexity of the composition, it was necessary to give an additional score for what could be regarded as perhaps the most important aspect of a communication in a second language, that is its intelligibility.

It was decided to give relatively little weight to the factor of grammatical and/or lexical accuracy, since measures of accuracy, unless made in the form of highly unreliable global judgements, tend to involve rather lengthy procedures, in that they require attention to specific errors. Also, the counting of errors, even if related to the number of words produced, can create a distorted picture of the child's writing ability, since more adventurous children may well attempt more complex structures, and therefore make more mistakes, without necessarily producing a less effective communication.

For the NFER tests, a scale of general intelligibility, inevitably subjective to some degree, was applied to each T-unit in turn, the scale being as follows:

0—completely unintelligible;
1—partially intelligible;
2—completely intelligible.

A bonus mark yielding a maximum score of 3 was given for any T-unit that was completely accurate, in all respects except that of minor spelling mistakes, these being ignored. A minor spelling mistake was defined as one which 'did not interfere with intelligibility'.

An example of a scored piece of writing is given below:

I like to come to^{3} school/ and my^{2} mummy like
school/~~so like wanet I saie to come school~~ and my
bearther we was so happe^{1} wane I'come home/
I'eat cake^{1} and daere my tea/ I halp^{3} my mummy./
One day I wanet in^{2} m sister house/ the end.

Intelligibility Score = 12
Mean T-unit length = 6·8

In the above example, the spelling mistake in the T-unit, 'I halp my mummy' is regarded as a minor one. However, 'Daere' to mean 'drink' in the preceding T-unit could be said to be a major spelling error which interferes with intelligibility.

The result of the application of the above system of scoring was that each essay received two separate marks, one for complexity (the average T-unit length) and one for intelligibility (the sum of the intelligibility score for each T-unit).

This system proved simple to administer, yielded relatively high inter-rater reliability (0·98 for the total score of the essay), and could be said to deal effectively with those aspects of a piece of writing which reflect most clearly the development of writing skill in a second language. The use of the system could not, however, justifiably be extended to deal with the scoring of samples of writing produced by adults, in their native language. Here, factors of style could well operate to produce relatively short T-units which would in no way indicate poor writing ability.

The performance of Asian children on the writing tests

In the final stage of test development, the tests were administered to samples of Asian and West Indian children with up to five-and-a-half years schooling in Britain. It was felt that, for the purpose of analysing the results of these second language learners, it would be more appropriate to divide the sample on the basis of length of stay in Britain, rather than on the basis of age.

For the Asian sample, differences in mean T-unit length were observed: the longer the children had been in this country, the longer was their mean T-unit length. The same results were not found in the case of the West Indian sample, who perhaps could not be viewed as true second language learners and whose results might be better analysed by age, as would be done with native speakers, rather than by length of schooling in this country.

For Asians with up to one-and-a-half years schooling in the UK (numbering 228) the mean T-unit length was 7·70, for those with one-and-a-half to three-and-a-half years schooling (numbering 281) it was 7·95, and for those with three-and-a-half to five-and-a-half years schooling (numbering 242) it was 8·71. These three length-of-stay groups are hereafter termed LOS 1, LOS 2, and LOS 3 respectively.

The rise in mean T-unit length from the short stay to the middle stay group is not statistically significant, but the rise from the middle to long-stay group is significant at the 99·9 per cent level.

From these results it seems that the T-unit can be regarded as a useful index of complexity of structure in the written work of second language learners, and that with Asian children, a slight increase in complexity of structure and/or clause length is observed from the one-and-a-half to three-and-a-half year stage, and a more definite one in the three-and-a-half to five-and-a-half year stage.

The generally greater improvement in writing skills shown by Asians after three-and-a-half years at school in Britain than before this point is borne out by the results on the other writing tests; this suggests that the greater increase is score on the essay task during the later period in a reflection of a real improvement in proficiency, rather than merely an artifact of the scoring system.

The essay task formed the last of a series of writing tests of increasing difficulty. The other tests in the series were as follows:

Writing One:	Sixteen pictures of objects and nine pictures of activities were presented to the child, whose task was to write beneath each picture the name of the object or activity.
Writing Two part one:	Here the task was to write an appropriate phrase or sentence beneath each of ten pictures.
Writing Two part two:	Ten sentence beginnings were provided which the child had to complete.
Writing Three part one:	The task was the same as that in the previous test, but here the sentence beginnings demanded the production of a more complex response, consisting of a minimum of two clauses.

Scores on the two parts of Writing Two correlated extremely highly and so were combined for the purposes of the analysis of results.

The mean scores for Asians in the three length-of-stay groups were compared on Writing One, Writing Two, and Writing Three part one. The difference in score between LOS 2 and LOS 3 was significant in the case of all three tests, as was the difference in score between LOS 1 and LOS 2 for all tests except Writing One. However, on all three tests, the difference in score was markedly greater between LOS 2 and 3 than between LOS 1 and 2. These tests were scored on the basis of general intelligibility, no account being taken of structural aspects of the written response. It would seem therefore that it is the general capacity of Asian children to write intelligibly in English and not merely their ability to produce more complex structures, which shows greater improvement after three-and-a-half years in this country than in the preceding period.

The speaking tests

The speaking tests, all of which correspond to the writing tests (except for one addition which aims to elicit direct speech) are scored by the same methods—using a general scale of intelligibility and dividing up the open-ended responses by means of the T-unit.

The speaking task equivalent of the written essay consists of a description of a series of three pictures. The response is recorded on tape, and written down by the marker to be scored according to the procedure used for the written essay.

It was decided that the T-unit might be a valid means of dividing up speech (as well as writing) as it had already been used for this purpose with some success by R. O'Donnell using English speaking children.

In 1967, O'Donnell *et al.* studied 1,000-word samples of the speech of 30 students in each of six age-groups, and found a similar steady increase in T-unit length with age, indicating that this measure of linguistic maturity can be applied equally successfully in both of the productive language skills.

With the Asian children, an increase was found in the mean T-unit length with increased length of stay as was found with writing. The children who had been at school in the UK for up to 18 months had a mean spoken T-unit length of 5·45 words, those who had been in the UK for between one-and-a-half years and three-and-a-half years had a mean spoken T-unit length of 6·09 words and those who had been in the UK from three-and-a-half to five-and-a-half years had a mean spoken T-unit length of 6·43 words.

However, as the speaking tests were administered only to a sub-sample of one in ten, the numbers tested are very small—20, 26 and 27 respectively—and so little weight can be given to these results. They do suggest, however, that the T-unit can be used in the scoring of spoken work as well as in the scoring of written work; and that it is possible that the mean spoken T-unit increases with length of exposure to the second language.

The above findings are of an essentially tentative nature, having arisen out of the process of test development. A full-scale survey was, however, conducted in April 1973, from which it should be possible to draw firmer conclusions regarding the proficiency in English of various immigrant groups.

References

HUNT, K. W. (1965). *Grammatical Structure Written at 3 Grade Levels.* Champaign, Illinois: National Council of Teachers of English.

O'DONNELL, R. C. *et al.* (1967). *Syntax of Kindergarten and Elementary School Children.* Champaign, Illinois: National Council of Teachers of English.

British Journal of Educational Psychology

Editor: J. D. NISBET, Department of Education, University of Aberdeen, Scotland

Vol. 44 Part 1 February 1974

The *British Journal of Educational Psychology* is concerned mainly with reporting the results of experimentally-based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry. The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words). There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists.

W. YULE, *et al.*: Over- and Under-Achievement in Reading: Distribution in the General Population.
B. BRYANT and P. E. TROWER: Social Difficulty in a Student Sample.
N. CHERRY: Components of Occupational Interest.
C. CRANE: Attitudes Towards Acceptance of Self and Others and Adjustment to Teaching.
M. ANNETT and A. TURNER: Laterality and the growth of Intellectual Abilities.
F. H. WALLBROWN, J. BLAHA and R. J. WHERRY: The Hierarchical Factor Structure of the Wechsler Adult Intelligence Scale.
S. A. BROWN: Scottish Science Teachers' Perceptions of Effective Science Teaching.
W. A. REID and B. J. HOLLEY: The Factor Structure of Teacher Attitudes to Sixth Form Education.

RESEARCH NOTES:
T. HONESS and P. KLINE: Extraversion, Neuroticism and Academic Attainment in Uganda.
D. DUCKWORTH and N. J. ENTWISTLE: Attitudes to School Subjects: a Repertory Grid Technique.
D. J. HARGREAVES: Situational Influences on Divergent Thinking.
C. E. CHANNON: The Effect of Regime on Divergent Thinking Scores.

BOOK REVIEWS

PUBLICATIONS RECEIVED

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Creativity and Curriculum Structure

E. Ogilvie, Northampton College of Education

Summary

It is contended that present assumptions regarding the relationship of creativity to curriculum structure are inadequately supported by both experimental findings and theoretical rationale. There are two main issues requiring further elucidation: the first concerns the influence of school environment on creativity directly: the second comprises an examination of the interrelationships of creativity, intelligence and curriculum structure.

The evidence presented demonstrates that school environments have a perceptible influence upon creativity but the generally held assumption that progressive environments are more favourable is shown to require careful qualification. The findings also suggest that the interrelationships of creativity, intelligence and curriculum structure are highly complex. Structure does not seem to affect creativity-intelligence correlations to any great extent: the major influences here would appear to be largely dependent upon the average intelligence of the groups and the extent to which the creativity tasks require school learning and especially verbal skills for their performance.

Introduction

A consideration of the literature devoted to discussions of the relationship between creativity and educational environment reveals a situation which is in need of much closer examination than it has received hitherto. In the first place, there is general presumption that 'informal' or 'progressive' schools provide an atmosphere conducive to the development of creative potential, and that the 'formal' or 'traditional' environment is highly inimical to the creative child (see Torrance, 1962 p. 70f). It is difficult to ascertain how far terms like 'progressive', 'traditional' and 'creative' indicate relatively simple and generally accepted concepts, but we refer here to those schools which emphasize so-called creative activities, self-initiated learning and discovery on the one hand, and to schools which concentrate on the didactic and authoritarian approach on the other.[1] The attributes of the creative child are listed by Torrance (1962) and described by traits in Guilford (1959). Basically the creative child is a nonconformist who questions authority and may even hold conformists in some disdain. He is an individualist with a proclivity for experimentation.

There are no agreed theoretical bases to support the common view that traditional environments are detrimental to creativity. It may be that its origins lie in accounts of the unhappy childhood experiences of a number of highly creative individuals.[2] Some firm evidence is provided by Haddon and Lytton (1968) who found 11-year-old English children to be significantly superior in divergent thinking abilities when taught in 'informal' schools. These authors cite other sources, Torrance (1962), Vernon (1964), Spears and Hilgard (1964), but point out that there is really very little evidence which directly tests and supports the hypothesis that 'informal' schools provide the best climate for creativity. Guilford (1967) offers something of a theoretical background supporting the general viewpoint but his remarks are based largely on generalities. Stoddard (1959) states the dogma lucidly but again without a single reference to hard evidence.[3]

Torrance's view is clearly that an informal atmosphere fosters creativity since, on the basis of wide experience of creativity testing, he cites approvingly (1964) the comment 'rigid authority deprecates ideas offered by pupils'. It is interesting, therefore, to notice that Torrance also believes the relationship between intelligence and creativity depends to no small extent upon the degree of structuring displayed by the school curriculum. Thus he seeks to explain Hasan and Butcher's (1966) failure to differentiate intelligence from creativity by arguing that the Scottish secondary school environment is too formal to permit the true facts to reveal themselves and that it brings about an artificially high correlation between creativity and intelligence quotients.

---

[1] Perhaps the best description of the dichotomy is in Barker Lunn (1970), p. 15 and ch. 3.

[2] Goertzel, M. & V. (1965) and Burt, C. (1970).

[3] See Hildreth, G. (1966), p. 470, 'there is no clear proof that informal classroom atmospheres are more likely to develop creativity of a high order than traditional curriculum organization and teaching practices'.

Dacey, Madaus and Allen (1969) however report only very low IQ/CQ correlations in their samples of Irish children, and it seems unlikely that the Irish secondary grammar schools were significantly less traditional than their counterparts in Scotland.[1]

Nevertheless, these authors by way of explanation for their low intercorrelations also appeal to factors linked to curriculum structure, and suggest that higher creativity-intelligence intercorrelation may be the result of 'an interaction between method factors relating to both measures rather than to a high relationship between the two traits'.

Haddon and Lytton (op. cit.) specifically predict higher correlations between creativity and intelligence in the informal schools; a view diametrically opposed to that given by Torrance. Haddon states that in the informal schools, 'because of their particular approach to learning, the link between the two aspects would be maintained to a greater extent' and continues, 'the idea is more forcibly expressed by saying that formal education will tend to destroy this connection by putting a premium on convergent thinking and conformist behaviour'.

It is noteworthy that none of the authors cited appears to be very clear about the nature of the 'link' between intelligence and creativity. Indeed the explanation offered amounts to little more than a restatement of the facts as they appear in the particular research findings; and there can be little doubt that the theoretical interpretation of interrelationships of creativity, intelligence and curriculum structure is unsatisfactory.

The present situation thus presents certain difficulties with regard to experimental exploration in the field. It is usual to match schools on some variable or variables before comparing their performances on creativity tests. Matching is necessary when (a) the matching variable correlates significantly with the criterion variable, (b) the criterion variable is in some real sense dependent upon the capability measured by the matching variable and (c) the independent variable and the matching variable are unrelated. Intelligence is frequently seen as meeting these conditions, and matching groups on verbal reasoning tests has become almost standard procedure in connection with research on teaching methods and related problems.

The works of Vernon (1964) and Pidgeon (1970) demonstrate that IQ is no longer thought of as an unchanging, innate attribute. Many of the experiments they cite show clearly that there is little warrant for treating intelligence as the variable independent of curriculum structure. As Dacey *et al.* (op. cit.) suggested, both intelligence and creativity can be considered as variables which are likely to be appreciably dependent upon the teaching methods employed in schools. Moreover, it is not unreasonable to guess that very informal schools are unlikely to provide environments conducive to the production of high Verbal Reasoning Quotients whereas formal schools may frequently approach optimum conditions in this regard (Wiseman, 1966; Butcher, 1968). If this latter hypothesis is true[2] and, in addition, a positive correlation is found between intelligence and creativity, then it appears that to match formal and informal schools on verbal reasoning is effectively to favour the informal school in any comparisons on creativity. Where the informal environment favours creativity the apparent effect will be exaggerated, and where the formal school favours creativity the effect will tend to be masked. Haddon and Lytton take no account of this possibility. Indeed a special point is made (p. 176) of the fact that informal schools have a mean VRQ slightly, but not significantly, below that of the formal schools, yet, in spite of this, the Informal Schools achieve results which are significantly superior'.

The conclusion we draw so far is that, in the absence of other information, there is no point in matching schools on VR tests in research into the effects of curriculum structure and that this procedure might be particularly misleading where a very wide range of environments is deliberately built into the research design.[1]

---

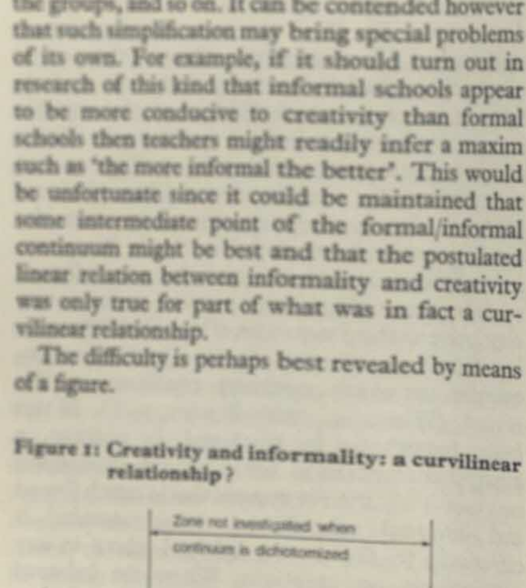
[1] *Times Educ. Supp.* 2.6.72, p. 10; Dacey's schools were staffed by religious orders.

[2] Clear supporting evidence is given in the Surrey Education Research Association's Report (1968), *To Stream or Not to Stream.*

This point introduces a second consideration which deserves attention. Researchers often content themselves with an exploration of the informal-formal dimension in terms only of a dichotomy. This is consistent with the form in which the problem is usually posed, and may be justified on the grounds of simplification, larger numbers in the groups, and so on. It can be contended however that such simplification may bring special problems of its own. For example, if it should turn out in research of this kind that informal schools appear to be more conducive to creativity than formal schools then teachers might readily infer a maxim such as 'the more informal the better'. This would be unfortunate since it could be maintained that some intermediate point of the formal/informal continuum might be best and that the postulated linear relation between informality and creativity was only true for part of what was in fact a curvilinear relationship.

The difficulty is perhaps best revealed by means of a figure.

Figure 1: Creativity and informality: a curvilinear relationship?



In the following sections a theory of creativity is outlined and relevant experimental findings are considered in the light of the foregoing discussions.

A theory of creativity

Research into creativity and curriculum structure must be based upon a theory of creativity which is made quite explicit, and the theoretical standpoint adopted in the present research derives from the Associationist Theory of Creativity as outlined by Mednick (1962). We postulate that creativity depends firstly upon the size and quality of the association reservoir. Mednick argues that an individual without the requisite elements in his response repertoire can hardly be expected to combine them in a creative solution: a child who has not learned his numbers will never combine them in novel number series. He further suggests that the greater the number of associations which a person has constructed amongst the requisite elements, the greater the probability of discovering new combinations.

Creativity depends secondly upon the relationship between the creative personality and the curriculum structure which helps to shape that personality. The general view of the effect of school environment on creativity noted earlier takes account only of this second point since it stresses simply the deleterious and inhibitive effects of authoritarianism and conformity pressures. It neglects or ignores the fact that a very 'progressive' or unstructured school situation might fail to teach children those very elements of learning and knowledge which are essential to the solution of problems included in creativity tasks.

It can reasonably be argued therefore that, beyond a certain point, a curriculum becomes so unstructured that it fails to provide a vital condition for creativity beyond mere freedom, namely, that of a basic achievement level or association reservoir. Moreover, it seems highly probable that some creative children are inhibited in the very informal environment simply because, in its extreme forms, it fails to reinforce creative behaviour consistently. Some 'progressive' teachers certainly eschew any form of extrinsic reinforcement such as the award of stars, marks, and verbal praises,[2] despite the fact that there is no evidence to show that creative behaviour is responsive only, or even especially, to intrinsic reward systems.[1]

---

[1] The problem is discussed in Barker Lunn (op. cit.), p. 58, where matching on VR tests was rejected in favour of matching on a Reading Test for reasons similar to those given here.

[2] See, for example, Brown, M. and Precious, N. (1968), p. 39; and Ridgway, L. and Lawton, I. (1968), p. 134.

Given the theoretical basis outlined above, it is possible now to summarize the relationships between creativity and curriculum structure as follows:

(1) Extreme environments of both 'formal and 'informal' types will be relatively inimical to creative expression. The highly formal environment will fail to provide opportunities for non-conformity, and the highly informal setting will be unfavourable to the development of sizeable association reservoirs. Both environments will fail to reinforce creative behaviour in any consistent manner.

(2) The environment most favourable to creativity will be that which provides for both freedom of expression and good quality association reservoirs. It will be especially favourable if extrinsic rewards are attached to creative behaviours.

(3) An environment so lacking in structure as to approach a state of anomie will be worst of all since it provides neither freedom nor basic achievement skills. It quite fails moreover to provide any reinforcement outside a random schedule.

(4) Only very extreme environments are likely to be totally hostile to creative output since some creativity tasks require very little in the way of school achievement and verbal learning. At more intermediate levels formal environments might provide relatively congenial conditions for success on tests requiring relatively high levels in basic skills. Informal curricula on the other hand might be relatively more effective in areas where the basic achievement levels required are relatively low. It is convenient at this point to note that in any thoroughgoing investigation it would be important to examine the exact nature of the differences between the various intermediate grades of school climate in terms of the individual creativity tasks and the minimum levels of achievement necessary to their performance.

The Associationist Theory predicts a positive correlation between intelligence and creativity, regardless of curriculum structure and supports a 'threshold' theory with regard to this interrelationship. It suggests that the influence of structure on CQ/IQ intercorrelations is likely to be highly complex, and warns against easy generalization in this regard.

Details of the investigation

The investigation below sheds some light on these theoretical problems though it was not designed specifically to answer them. The findings are therefore not without interest in the context of the ideas which have been put forward, and they constitute as least a partial excursion into the field. The essential experimental details are summarized below:

Creativity tasks

The tests used owe much to Guilford's and Wallach and Kogan's (1965) conceptions with regard to both type and task adminstration. Taken together they are measures of complexity, associative fluency and originality and were given in situations free from time and examination stresses. Game-like approaches based on the establishment of friendly rapport with the children were invariably used. Five 'games' were involved.

1. The Uses Task. Children are asked to write down all the possible uses for a newspaper, shoe, button, chair, knife, tyre, cork and key. An example of an unusual response for a 'button' would be, 'raise table leg' as compared with, 'fasten coat'.

2. The Line Meanings Task. This procedure involves visual stimulus materials whereby the child is confronted with eight different line diagrams and asked to generate meanings relevant to the line in question.

3. The Number Series Task. This instrument is a form of the familiar Make Up Problems Tests (Getzels and Jackson, 1963) and requires the child to construct a set of different mathematical series. It is similar to Westcott's (1968) assignments and is scored for both fluency and originality.

The contrast between commonplace and original responses can be seen in the examples given below:

(i) 2, 4, 6, 8, 10 . . .
(ii) 15, 13, 11, 9, 7 . . .
(iii) 6, 7, 12, 13, 24, 25, 48 . . .
(iv) 50, 47, 44, 42, 40, 43, 39, 36, 33, 31, 29

4. The Symbol Series Task. This instrument is similar to Task 3 but symbols and/or letters are used instead of numbers. It is scored for fluency only and an example would be: □○ △ □○△□.

5. The Kites-Fishes Puzzle Design Task. Children are here required to construct a puzzle commonly found in children's puzzle books.

---

[1] See DIENES, Z. P. (1964), ch. 8, for a discussion of ntrinsic and extrinsic motivation.

Three children are envisaged as flying three kites or fishing for three fishes and the child must draw in the three tangled strings so as to create the most complex puzzle he can. He is further required to design original kites and fishes so that complexity and originality scores can be constructed.

The Schools

Five schools well known to the author, several colleagues and an LEA administrator were selected and ordered on the 'structural' dimension as follows:

School	*Type of Environment*
A	Highly structured (Formal)
B	Midroad I
C	Midroad II
D	Unstructured (Informal)
E	Indeterminate (Anomie)

The main criteria used in identifying the 'structure' dimensions are indicated in Table 3.1 of Barker Lunn (op. cit.), p. 27.

School E is described as in a state of anomie since the general organization lacked coherence and it presented features of both types of environment always attenuated by an air of casualness.

The effects of differential home backgrounds are confounded in the findings since school E contained wholly working class children. The other schools were more comparably homogeneous in terms of parental social status.

The subjects

The children comprised the fourth-year groups (10- to 11-year-olds) in each school. The numbers varied somewhat between the tasks owing to the fact that testing sessions had to be spread over the whole summer term. The school B group was tested one year later than the other groups for practical reasons irrelevant here. The total number involved ranged from 131 to 161 children.

Results and discussions

The results are set out in the following tables.

The strength of the highly significant relationship amongst creativity scores (Table 1) is indicated by Haggard's (1958) R=·20 or the correlation ratio E=·44. These results are thus consistent with the main hypotheses put forward earlier in that neither of the less structured environments produce children with higher scores on the creativity tests and the most uncertain classroom organization gives rise to the lowest scores.

The question as to how far these findings conflict with those of other workers is difficult. The fact that our extremely undirected school E seems least conducive to the creative development of children can readily be reconciled with the relative success of Haddon and Lytton's informal group (1968) by postulating that our midroad environments are akin to their informal type. Their school D seems in fact to resemble our school C quite well, and the degree to which our school D is fairly extreme may be judged from the headteacher's remark, 'We teach our pupils not to be able to do intelligence tests'. This explanation of the discrepancy between present findings and those of Haddon and Lytton is attractive since alternative proposals such as incomparability of tests or testing situations do not seem to be particularly impressive. It is important therefore to stress that acceptance of the view put forward here implies that qualifications to Haddon and Lytton's major thesis are very important. The proposal that 'informal progressive teaching would promote these (divergent thinking) abilities more than formal subject-centred teaching' can at best be only partially true, and could easily be a dangerous

Table 1: Comparison by schools of the mean creativity aggregate scores: summary data and analysis of variance

SCHOOLS	A *Structured*	B *Midroad I*	C *Midroad II*	D *Unstructured*	E *Anomie*
Means	206	231	227	199	171
Source	s.s.	df	ms	F ratio	
Between	53,292	4	13,323	7·54 (·01)	
Within	222,702	126	1,767		
Total	275,994	130			

Table 2: Intercorrelation of VHQ and CQ in terms of curriculum structure

Curriculum Type	*Structured*	*Midroad I*	*Midroad II*	*Unstructured*	*Anomie*	*Total Group*
Correlation rho	·49**	·17	·41*	·31	·17	54·**
N	31	41	20	20	19	131
Mean CQ	101·1	115·1	109·5	99·6	96·2	106·1

**<·01 *<·05

half-truth at that. It has been made clear here that the 'structure' dimension involved both requires and is susceptible of a greater measure of subdivision than a simple dichotomy provides. The main reason for greater subdivision is that the relation of creativity to curriculum structure is in all probability much more complex than it appears and oversimplification and generalisation can only do more harm than good. This is particularly true in an area where many educators appear to have already formed opinions in the absence of empirical evidence and seem willing therefore to fasten on any research findings which appear to support their views (Hudson, 1970).

The evidence of the present research to some extent goes against popular opinion. It shows the formal schools tested were less inimical to creative potential and expression than were the types of progressive schools tested, and that this may be more true of some curriculum areas than others. There can be little doubt that the Midroad environments were best in general terms, but we suggested earlier that this might be due merely to their failure to suppress rather than to their making any really positive contribution to the development of creative potential. Further research in this area ought to specify more closely than has been possible here, the exact nature and range of the environments involved, using perhaps the checklist given by Barker Lunn as a basis for more objective categorizations. It ought to investigate individual differences in response to differing environments in order to fill in some of the theoretical gaps to which attention has been drawn.

The evidence with regard to the interrelationships of intelligence and creativity as reported in Table 2 suggests that different types of children are differentially creative in different environments.

The evidence unfortunately is meagre, even confusing, and the main value of the table is that it underlines the need for further research oriented towards solving the theoretical issues raised earlier. Depending upon which group is chosen, either Torrance's or Haddon's position receives some support; and yet both are contradicted. It seems that in all probability the interrelationship between VRQ and CQ depends primarily upon the general intelligence level of the group concerned and only to a relatively slight extent on curriculum structure as such. There is now a growing body of research which amply supports the 'threshold' concept of this relationship, (Haddon, op. cit., Getzels and Jackson, op. cit., Hudson, 1966) and Wallach and Kogan's claim to have discovered an independent dimension can be discounted on the grounds that their group averaged c. 120 IQ. The present results given in Table 3 display the 'threshold' phenomenon quite nicely and are thus in line with the 'Associationist' Theory of Creativity. They suggest that the

Table 3: Intercorrelations of creativity and intelligence

		VRQ		Correlation	
	Range	Mean	SD	CQ VRQ	*N*
Creativity:	Total	105·8	15·2	·547	131
	Upper Half	112·8	12·2	·133	66
	Lower Half	99·0	14·7	·537	65
Intelligence:	70–139	105·85	15·18	·547	131
	107–139	117·99	8·40	·108	66
	70–107	93·90	10·00	·484	65

'threshold' is, however, well below the 120 IQ level often proposed (McKinnon, 1962).

Conclusion

The relationship of creativity to curriculum structure is more complex than has been admitted in previous work. The commonly held opinion that 'informal' environments are generally more congenial to creative expression requires careful qualification and in our present state of knowledge any generalization is dangerous. It is suggested that until individual differences in creative response to various degrees of curriculum structure are more fully investigated the most conducive environment is that characterized here as 'midroad'.

The 'threshold' theory of the interrelationship between intelligence and creativity receives considerable support from present findings which are wholly in accord with those of Haddon and Lytton at this point. The threshold is likely, however, to be well below the 120 IQ level sometimes suggested. The Associationist Theory carries no clear implications for the interrelationships of creativity, intelligence and curriculum structure and is in need of extension. Meanwhile, there is little warrant for attempting to explain conflicting findings with regard to creativity-intelligence correlations in terms of differences in curriculum structure. A far more plausible explanation would seem to derive from the concept of an intelligence threshold.

References

Barker Lunn, J. C. (1970). *Streaming in the Primary School.* Slough: NFER.

Brown, M. and Precious, N. (1968). *The Integrated Day in the Primary School.* London: Ward Lock Educational.

Burt, Sir C. (1970). Foreword to Kelmer Pringle, M. L. *Able Misfits.* London: Longmans.

Butcher, H. J. (1968). *Human Intelligence.* London: Methuen.

Dacey, J., Maudus, G. and Allen A. (1969). 'The relationship of creativity and intelligence in Irish adolescents', *Brit. J. Educ. Psychol.*, 39, 261–7.

Dienes, Z. P. (1964). *Mathematics in the Primary School.* London: Macmillan.

Getzels, J. W. and Jackson, P. W. (1963). *Creativity and Intelligence.* New York: Wiley.

Goertzel, M. and V. (1965). *Cradles of Eminence.* London: Constable.

Guilford, J. P. (1959). 'Traits of creativity', in Anderson, H. (Ed.) *Creativity and its Cultivation.* New York: Harper & Row.

Guilford, J. P. (1967). *The Nature of Human Intelligence.* London: McGraw Hill.

Haddon, P. A. and Lytton, H. (1968). 'Teaching approach and the development of divergent thinking abilities in primary schools', *Brit. J. Educ. Psychol.*, 38, 171–80.

Haggard, E. A. (1958). *Interclass Correlation and the Analysis of Variance.* New York: Holt, Rinehart & Winston.

Hasan, P. and Butcher, H. J. (1966). 'A partial replication of Getzels and Jackson's study', *Brit. J. Psychol.*, 57, 129–35.

Hildreth, G. (1966). *Introduction to the Gifted.* New York: McGraw Hill.

Hudson, L. (1966). *Contrary Imaginations.* London: Methuen.

Hudson, L. (1970). As reported in *Times Educ. Suppl.*, 24.7.70, p. 9.

Mednick, S. A. (1962). 'The associative basis of the creative process', *Psychol. Rev.*, 69, 3.

McKinnon, D. W. (1962). 'Nature and nurture of creative talent', *Amer. Psych.*, 17, 487–8.

Pidgeon, D. A. (1970). *Expectation and Pupil Performance.* Slough: NFER.

Ridgway, L. and Lawton, I. (1968). *Family Grouping in the Primary School.* London: Ward Lock Educational.

Spears, P. S. and Hilgard, E. R. (1964). 'The teacher's role in the motivation of the learner', in *Theories of Learning*, 63rd NSSE Yearbook, Univ. of Chicago.

Stoddard, G. D. (1959). 'Creativity in education', in Anderson, H. (Ed.) *Creativity and its Cultivation.* New York: Harper & Row.

Surrey Education Committee (1968). Report of the Research Association, *To Stream or not to Stream.*

Torrance, E. P. (1962). *Guiding Creative Talent.* Englewood Cliffs. New Jersey: Prentice Hall.

Torrance, E. P. (1964). *Role of Evaluation in Creative Thinking.* University of Minnesota.

Vernon, P. E. (1964). 'Creativity and intelligence', *Educ. Res.*, 6, 163–9.

Wallach, L. M. and Kogan, H. (1965). *Modes of Thinking in Young Children.* New York: Holt, Rinehart & Winston.

Westcott, M. R. (1968). *The Psychology of Intuition.* New York: Holt, Rinehart & Winston.

Wiseman, S. (1966). 'Environmental and innate factors and educational attainment', in Mead, J. E. and Parkes, A. S. (Eds.) *Genetic and Environmental Factors in Human Ability*, 65–80. Edinburgh: Oliver & Boyd.

Linked Courses in Herefordshire

Dennis H. O'Donnell, Queen Elizabeth School, Bromyard, Herefordshire

Summary

The enquiry outlined below is part of one which surveyed the school-college linked course in the County of Herefordshire. As a result of the survey it was possible to construct bases for the judgement of the relative success of the linked course as a part of the secondary curriculum.

Comments from seventy teachers and others engaged in education together with statements from the relevant literature on the subject were used to write an attitude scale. This scale was applied to 273 past and present pupils of a mixed four-form entry rural comprehensive school with about 550 pupils. The measure of the success of linked courses was derived from the results of the test of the pupil's attitude.

The development of linked courses

Linked courses are used to provide an element of vocational education in the school curriculum. 'Vocational education' is here used to mean learning about the world of work, learning elementary skills and appreciating their place in the general scheme of things related to the particular occupation in which the skill is used. Vocational studies aim to be relevant to the pupil's likely job choice when he leaves school and it is hoped would give him a degree of experience of as well as assistance in understanding the context within which he will work. It should help him to make a rational choice of occupation.

School-college linked courses have developed in response to the demand for vocational education. Further education colleges are clearly much better placed to provide the necessary teaching and equipment than most moderate size secondary schools. In many areas secondary school pupils attend colleges on a half or whole day a week on such courses. Fourteen thousand students in Britain were on linked courses in 1969 (Brown, 1971). In the fourth year of his schooling a boy might for example spend his time learning elementary building skills (brickwork, carpentry, plastering, plumbing, etc.), elementary motor mechanics (how to deal with simple faults, how the internal combustion engine works, safety in the workshop), catering (production kitchen work, baking, waiting), whilst a girl might be taught typewriting with elementary office skills and accounts or she might follow a course in child-care with elements of pre-nursing training (baby and toddler care with institution visits and human physiology and hygiene). Courses of a wide variety are described by Brown (1971), Skinner (1968 and 1971), O'Donnell (1966 and 1973), and Davies (1968). The courses include hairdressing, engineering, window-dressing, printing and agriculture. In some colleges 'taster' courses giving a wide variety of learning experiences are offered for short periods. No attempt is made to turn out a craftsman or a fully trained boy or girl. The experience is the important thing. Schools and colleges have joint staff meetings, where this is appropriate, to link the work done in both the school and the college.

The Newsom Report (CACE, 1963) and Schools Council Enquiry I (1968) both indicated a desire on the part of pupils and their parents for more vocational studies in school. Whilst the larger secondary school will be better placed to satisfy such demands the smaller school will

---

Acknowledgements

The author wishes to acknowledge the great help and encouragement given to him by Professor Jac L. Williams, Dean of the Faculty of Education, University College of Wales, Aberystwyth, and his thanks go to Mr A. S. Ghuman of The Education Department of the same College for his invaluable help in the design of the original enquiry and in the interpretation of the results.

tend to look to the nearby college for assistance. The likely economies of scale and the most economical use of resources by a number of schools making use of the existing facilities is referred to in the Newsom Report. Davies (1968) describes a framework by which schools and colleges may be brought together to produce the minimum of friction and the maximum value for the student on linked courses.

In the school that is the subject of the investigation outlined here, courses offered to fourth- and fifth-year pupils were run in such a way that linked courses as alternatives were offered in place of the more difficult specialist sciences, a second foreign language and minority subjects for examination such as RE or music. These alternative courses were offered in the fourth and fifth forms to a comprehensive population of about a hundred and ten pupils in each year group. They were

(i) a building course leading to a CSE in building construction;

(ii) motor vehicle maintenance leading to a CSE in the subject;

(iii) catering, leading to City and Guilds Certificates;

(iv) commercial subjects, typing, commerce and accounts leading to CSE examinations;

(v) child-care and pre-nursing, a composite course which combined a certificate in child-care and a CSE examination in human physiology and hygiene; and

(vi) a rural studies course shared between the schools and the County School of Agriculture where farm machinery and animal husbandry were specially dealt with.

Evaluation of linked courses as a curricular element

With the school leaving age now raised to 16, both colleges and schools are assessing anew the role of vocational studies in the curriculum. Schools will doubtless welcome additional facilities for their older pupils. They will also find that the type of course which can be offered to them will be of the kind that most appeals to those for whom provision has not in the past been wholly adequate. Colleges will face the withdrawal of 15- to 16-year-old day-release and full-time students in the year now in progress (1973-74) and if there is no sudden change in other age groups offering themselves for college courses, staff and facilities are going to be less under pressure than they are now. The Linked Course may therefore be seen as a device for helping schools and colleges alike.

Much re-thinking is going on at the moment in curriculum renewal and in curriculum development. The impetus for much of these developments doubtless springs from the raising of the leaving age which affected all schools in September 1973. For some heads of schools, educational goals of a vocational nature arising from the re-thinking going on may have little appeal. More have been persuaded to accept them by the pressure of opinion and the need to make meaningful educational provision for the new fifth-formers.

What follows is an attempt to achieve a degree of evaluation for the whole linked course idea. The enquiry as far as is known is the first attempt to tackle the problem in this way.

Pilot survey

One hundred and eighty pupils were asked to express their opinions about linked courses as an element in their curriculum. They were from nine different secondary schools in the county of Herefordshire and attended college on a day or a half day a week. This survey took place in 1970. It was clear that this 'day release' from school was popular. The boys and girls questioned were at the end of courses in commercial subjects, building, child care/pre-nursing and agriculture. Some of the answers to the questions put to these young people were revealing. Irrespective of the course followed most pupils agreed that they had been given a fair choice of courses when they began and said that they would choose the same course again if they had their time over again. A revision of ideas is however evident in the answers to these questions:

'Did you begin the course thinking that you would be helped by it in a particular job?'

(ii) 'Will you still go for that job?'

The percentage replying positively were:

	Commerce	*Building*	*CC/ Pre-Nursing*	*Agriculture*
(i)	90	76	78	52
(ii)	64	48	30	35

Clearly more young typists have their sights on a particular and related job than young farmers. All categories dropped in their determination to proceed. The reason might have been interesting to explore in every case. In any event it was clear that some thought it a successful venture even if no job was forthcoming. It was considered a try out and a part of general education. It was a change of scene with different people for a day a week. For the purposes of this enquiry it was decided to go for a measure of overall approval or disapproval without attempting to measure performances nor attempting to match the success rates in different courses.

Survey of staff opinion

This survey was to find the extent of the use of linked courses in the schools and to ascertain the feelings of the staffs of school and college as well as those of Careers Officers of the Youth Education Service was carried out by means of a detailed questionnaire. All the colleges and seventeen of the twenty secondary schools responded. Many of the statements for the attitude scale were derived from the survey responses.

The criteria by which success was to be judged and the attitude scale

From part of the survey questionnaire answers it was possible to compile a list of statements which categorized success of the linked course in the eyes of the staffs and careers officers. No attempt was made to give any one criterion more weight than any other. Broadly the criteria covered four main areas: the courses were judged successfully if (a) they met a pupil's need, (b) they improved his general behaviour and conduct, (c) they extended his horizons in deciding on a career and (d) they provided a general educational stimulus.

One hundred and nineteen statements were sent to 20 judges (heads, lecturers, employers and teachers). They were asked to put these statements into categories, based upon their judgement, that (in relation to the key statement) they were favourably or unfavourably orientated to. The key statement was: 'Linked courses form a satisfactory and successful element in the curriculum of the fourth- and fifth-year pupil in the secondary school'. To assist them in their judgement the word 'success' was clarified by giving them a detailed list of the criteria referred to above. The categories into which the statements were sorted were given scale values as:

Very, very favourable: 1, Very favourable: 2, Favourable: 3.
Slightly favourable: 4, Neutral: 5, Slightly unfavourable: 6.
Unfavourable: 7, Very unfavourable: 8, Very, very unfavourable: 9.

If a judge rated a statement about linked courses as one indicating that the courses were successful, he allocated this to the favourable category and vice versa. Thus one of the criteria of success is 'if a course clarifies career objectives'. Many judges in the light of this criterion thought that the statement: 'Linked courses cause boys and girls to enquire more about possible careers than they might otherwise do', showed a very favourable attitude to them. The statement was therefore given a scale value of two by these judges. Pupils agreeing with this statement were taken to be displaying a very favourable attitude and in terms of the criteria articulated by teachers, heads, lecturers, etc. the courses were judged to be very successful as a curricular element.

From the 119 statements, 35 were selected so as to represent the range between those which indicated success ('Linked Courses provide an excellent link between school and work') to those indicating failure ('Contact with more adult students undermines the pupil's attitude to school'). A cumulative frequency graph was drawn for each statement. Where the interquartile range was more than two points out of nine the statement was rejected. This rejection meant that there was lack of unanimity on the part of the judges. Thus the categories of the statements had to be reasonably firmly agreed by the judges. They had to be reasonably evenly spaced and evenly distributed between the categories into which they had been sorted. At the 'successful' end for example was the statement: 'Pupils who go on linked courses develop a much more responsible attitude to their work both at school and at college'—ranked 17 out of 35. A neutral statement was: 'A pupil who does badly at school is as likely to do so at a technical college on a linked course'—ranked 31. An unfavourable one chosen was: 'Linked courses make pupils commit themselves too early to a particular job'—ranked 24. The ranks indicated were those of the statements in the final scale with all responses totalled.

A further dimension was considered in selecting the items for the scale. Thus reasonable agreement on the scale value or category of the statement was a first requirement. Balance in the number of statements selected from each category came next and then an evenness over the whole range was looked for. The further dimension now considered was the significance of the subject matter of the statements. From the staff opinions obtained through the survey a group of topics relevant to the success or otherwise of the courses was teased out. These concerned: pupil motivation, conduct and general response to vocational work, careers and choice thereof, the world of work, the broad educational implications of such studies and the administrative problems thrown up by the scheme. Statements relevant to all these areas of concern were chosen. Some examples of a positive and negative kind that were used were: 'Because colleges are so different from schools, pupils tend to be hindered rather than helped by linked courses'—ranked 31; 'By getting started on something useful as on a linked course, boys and girls are encouraged to see beyond the skills they are introduced to'—ranked 8; 'Pupils who have had linked courses are less likely to change their jobs during the three years after leaving school'—ranked 21. From the pupils responses in the pilot survey one could identify areas of significance for them. Thus statements could be chosen which covered topics that they would have in mind. For example the effect of the courses in helping or hindering job choice was related to the statement: 'Linked Courses help pupils to make up their minds before they leave school about the kind of job they want to do'—ranked 6. A carrot of an unrealistic kind was implied in: 'Pupils are sometimes disappointed because they cannot get jobs related to the linked courses that they followed'—ranked 12. Whether pupils saw the course as a useful part of their general education was related to the statement: 'Linked Courses are useful in extending a young person's understanding of other people's lives and work'—ranked 3.

The scale when finally compiled was applied to three groups of people: two groups of pupils who had left and one group who were still at school. The first two groups were all the leavers

Table 1: Mean attitudes of all respondents

	1969–70 Leavers			1970–71 Leavers			1972 Pupils		
	n	*Score*	*sd*	*n*	*Score*	*sd*	*n*	*Score*	*sd*
Engineering (boys)	—	—		—	—		16	3·0	0·6
Commerce (girls) 4th Year	4	2·8	0·33	13	3·8	1·0	15	3·4	0·47
5th Year	11	2·7	0·47	11	3·0	0·50	12	3·2	0·30
Child Care (girls) 4th Year	4	3·5	0·33	9	3·3	0·65	14	3·6	0·66
5th Year	3	3·1	0·17	—	—	—	5	2·4	0·22
6th Year	—	—	—	2	3·0	0·65	2	3·0	0·16
Agriculture (RS) (boys) 4th Year	6	3·2	0·66	4	3·5	0·55	8	3·0	0·22
5th Year	—	—	—	—	—	—	3	2·9	0·22
Building (boys)	13	3·3	0·83	12	3·5	1·10	8	3·0	0·59
All boys and girls on Linked Courses	41	3·1	0·60	40	3·4	0·91	83	3·2	0·60
All boys not on Linked Courses	18	3·0	0·40	36	3·3	0·70	55	3·5	0·69
All girls on Linked Courses	22	2·9	0·40	24	3·4	0·80	48	3·3	0·64
All girls not on Linked Courses	8	2·9	0·55	14	3·3	0·91	21	3·1	0·68
All boys on Linked Courses	19	3·3	0·90	16	3·5	1·04	35	3·0	0·68
All boys not on Linked Courses	10	3·0	0·37	22	3·2	0·51	34	3·4	0·69
All boys	29	3·2	0·75	38	3·3	0·80	69	3·2	0·68
All girls	30	2·9	0·52	38	3·3	0·85	69	3·2	0·64
All boys and girls	59	3·1	0·48	76	3·3	0·83	138	3·3	0·67

	1969–70	1970–71	1972
Percentage of sample who responded	76	71	88

Scale Values: 1. Very, very favourable.
2. Very favourable.
3. Favourable.
4. Slightly favourable.
5. Neutral.
6. Slightly unfavourable.
7. Unfavourable.
8. Very unfavourable.
9. Very, very unfavourable.

All respondents: 3·26. sd 0·74.

in the years 1970 and 1971. The third group consisted of all the pupils in the fourth, fifth and sixth forms.

The results from pupils past and present

The mean attitudes of the respondents is given in Table 1 in the various categories into which they were divided. An overall mean attitude of 3·26 (sd 0·74) was recorded. The figure indicates a favourable attitude to linked courses as an element in the curriculum of the fourth and fifth-former in the secondary school. The pupils speak with an uncannily unanimous voice in their judgement. School and college staffs seem to have been correct in their assessment of what their pupils want.

The mean attitude for each year group was:

1970 leavers 3·1 (sd 0·48) i.e. Favourable,
1971 leavers 3·3 (sd 0·83) i.e. Favourable,
1972 pupils 3·3 (sd 0·67) i.e. Favourable.

No significant difference was found between these groups in their mean attitude. Some significant differences were found in the results of three groups of girls:

(i) between all girls and girls who were following the child-care/pre-nursing course in the fourth form in 1970. A degree of anti-school feeling which had perhaps 'spilled over' into their college work may have been to blame. This group had suffered some disturbance during the Comprehensive reorganization that had taken place at about this time.

(ii) between girls in the fourth form and those in the fifth and who were on the commerce course in 1971 and

(iii) between these same year groups on the child care/pre-nursing course in 1972.

A slight improvement in the attitudes to linked courses that they had followed on the part of fifth-formers could perhaps be explained by attributing to them a more mature attitude towards what they were doing as they experienced more of the course. Such a shift of attitude would certainly be encouraging if this is the reason. Schools have some apprehension that the reluctant and disinterested fifth-former will be difficult to motivate.

From an inspection of the rank order of the statements checked by responding pupils it was concluded that past and present pupils regarded linked courses as motivating forces and that they helped pupils to make up their minds about their careers. It seemed that pupils saw the courses as a significantly useful part of their general education and that they were not likely to be unduly disappointed if they did not get a job related to them. They did not see the courses as having an adverse effect on school discipline.

Of the total statements checked, the following were the most popular:

'A linked course gives a pupil who is not doing very well at school for some reason, a new start in new surroundings and with different people'.

'School courses ought to contain some vocational education—teaching about jobs and how to do the sort of jobs that skilled craftsmen do—in addition to ordinary subjects.'

'Linked courses are useful in extending a young person's understanding of other people's lives and work.'

The first has a message for us that suggests that an optimistic pupil should be allowed a second chance even though his current record may not be a very good one. One wonders who spoiled his first one. The second statement clearly tells what pupils want and the last one that they are not narrow-minded about vocational studies. Three other statements that were very popular with the various year groups were:

'Linked courses cause boys and girls to enquire more about possible careers than they might otherwise do.'

'Linked courses help pupils to make up their minds before they leave school about what kind of job they want to do.'

'On the whole linked courses satisfy a need for pupils in their last year or two of schooling.'

Again careers and the meeting of a sensed need are clearly in the mind of the pupil.

Conclusion

Eighty per cent of the pupils in the three year groups taken together responded to the enquiry. In doing so they provided much evidence in the letters that they wrote, that linked courses had been popular. One gave as his opinion that linked courses should 'take in more careers'. Other statements, such as:

'the course helped me to decide what career I wanted to take up'

'I found it helped me to decide what I wanted to do for a career'

'I attended a pre-nursing course and it helped me as I was going to take up nursing as a career' support the impression that at least in some cases, job clarification is helped by the courses. One boy felt that 'plumbing was a very satisfying occupation', an opinion he had apparently acquired on a linked course.

Very few of the pupils who wrote as well as sending back the questionnaires had critical comments. A few of them pointed out that their jobs on leaving were not related to their courses but on the whole did not think that this mattered. Their minds had been clarified and had been changed in some cases as a result of wider knowledge. More than one pointed out that 'many country school students do not realize the many careers open to them especially to those not keen on academic subjects', whereas on the other hand a plaintive cry came from one girl that 'it was all right for town girls to learn typing and office work' but she like others just could not get to town to follow her chosen job because the country bus service had been withdrawn.

One hopes that the evaluation attempted here may provide useful information in assessing the relative importance that schools might give to vocational studies linked with colleges. At this time heads and school staffs are very much aware of the need to diversify the curriculum in older year groups. The growth of the scheme in Herefordshire seems to have been one which has the approval of both the pupils and of those who teach them.

The strength of the attitudes held seems firm enough to reassure others who have doubts about introducing linked courses into their scheme of things. Some form of vocational work in the last few years of schooling is popular. An economical use of resources would be to achieve this by organizing linked courses.

References

Brown, S. (1971). 'School and college courses', *Trends in Education*, 23.

Central Advisory Council for Education, England (1963). *Half Our Future* (Newsom Report). London: HMSO.

Davies, J. L. (1968). 'Collaboration between secondary schools and technical colleges', *Vocational Aspects of Education*, 51.

O'Donnell, D. H. (1966). 'Newsom in rural England', *Times Educ. Suppl.*, 23.4.66.

O'Donnell, D. H. (1973). 'A survey and evaluation of school-college linked courses in the county of Herefordshire'. Unpublished MEd Thesis, U.C.W. Aberystwyth.

Schools Council (1968). *Young School Leavers. Enquiry I.* London: HMSO.

Skinner, W. E. (1968). 'School and college co-operation', *Trends in Education*, 11.

Skinner, W. E. (1971). 'A joint OND course', *Vocational Aspects of Education*, 23.

'O' Level or CSE?

The Consequences of Double Entry

Joan Fry, Rolle College, Exmouth

Summary

What kinds of results are attained when students are entered for both 'O' level and CSE in a subject? It is now increasingly the case that a school has candidates suitable for entering both types of examination. The study reported here looked at some of the consequences of double entry in 16 comprehensive schools which covered the complete ability range in one LEA. Seven academic subjects where double entry was most common are analysed in detail: in English, languages, history and biology, the number of additional 'O' level equivalent passes exceeded the number of cases where passes were obtained at both 'O' level and CSE; and in language and history over half the double entrants benefited because they failed 'O' level but obtained Grade One or Two in CSE. Conversely, the results show that if doubtful candidates entered only for CSE and not 'O' level a substantial number of 'successes' would be lost particularly in geography and to a lesser extent in English, mathematics and biology.

Introduction

The raising of the school leaving age, thereby doubling the numbers of fifth-formers in school, has made the questions surrounding examination at 16+ particularly important this year. The traditional pattern of 'O' level in the grammar school and CSE in the secondary modern was anyway becoming blurred in the comprehensive schools which are likely to have pupils suitable for entering both types of examination.

The relationship between the two systems of examination has been discussed frequently. The fairly general acceptance of CSE Grade One as 'O' level equivalent has probably helped to foster the hopes for a unified system but it has not solved all the problems for teachers deciding which examination is most suitable for their pupils. As in any selection process, it is the borderline cases which cause difficulties and the results of mistaken allocation can be far reaching. To leave at 16 as a 'failed "O" level' candidate may indicate to the initiated that the pupil showed promise at some stage, but he will be at a clear disadvantage compared with the holder of a respectable CSE certificate. For those hoping to enter higher education, however, good grades at 'O' level are necessary, and even a long list of CSE Grade Ones will be regarded with some caution.

The Joint GCE/CSE Committee of the Schools Council discussed the problem facing borderline candidates and their teachers in *Examining at 16+* (1966). Although the 'near miss' grade at 'O' level has been used unofficially by schools and pupils in making decisions about re-sits and choice of 'A' level studies, the Committee decided against formally acknowledging the grade, and recommended 'when in doubt use the CSE system' (para 69). To support the recommendation, they promised monitoring procedures to prevent the 'O' level standard being raised by the withdrawal of the weaker candidates. Any future decision to publish grades might encourage reliance on 'O' level for borderline candidates, but it might be several years before the value of a 'near miss' in the occupational or higher education markets could be determined.

The study: 'O' level or CSE?

In the present study[1] it has been possible to look at the examination practices of 16 comprehensive schools and the results obtained by their fifth forms (see also King, in press). The schools were within one local authority area and together provide for all those of secondary school age, apart from a few who go to the direct grant or church schools. Between them, therefore, they cover the complete ability range. The schools were entering their first comprehensive intake, but all had had experience of both examination

---

[1] This study is part of a larger project sponsored by the Social Science Research Council.

systems previously. It is hoped that schools in similar situations may find the information useful.

Studies of examination results often deal separately with 'O' level and CSE. It is inevitable that the various Examining Boards should report on their examinations in this way, but when looking at the performance of individual candidates or particular schools it is important to look at the two examinations together. Of the 1482 candidates entered by the 16 schools, 854 (58 per cent) were taking a mixture of 'O' level and CSE. In half of the schools, no pupils took entirely 'O' levels and only 51 (3·4 per cent) such candidates were found altogether in the other eight schools. These pupils had been skilfully selected, for all but six obtained five passes, usually with good grades. Indeed, the teachers seem to have been remarkably successful in determining candidates' potential, for only seven of the 577 candidates taking CSE only achieved five Grade Ones and might therefore have been successful on 'O' level courses.

But what happened to the majority of pupils, who were entered for a mixture of 'O' level and CSE? It seems clear that decisions about them were being made on an individual subject basis. In the more esoteric subjects, like Latin or cinematography, the school might have no choice and have to enter for 'O' level or CSE respectively, but in the more popular subjects the schools could shop around for the syllabus they preferred among the local CSE and several GCE Boards. In some subjects, a school entered all the candidates for the same examination, but in the traditionally important school subjects, pupils might be entered for either 'O' level or CSE, or both. While most teachers would accept the idea of entering pupils for a mixture of 'O' level and CSE in different subjects, double entries in the same subject are often frowned upon. The practice involves double entry fees and increases the proportion of the summer term spent in examinations. It smacks of mark grubbing rather than education and if the results are too discrepant it is bad for confidence in the validity of the examinations themselves. It is not surprising that there should be discrepancies in the borderline areas. Any examination, even if it includes an element of course work is merely sampling performance. Where the syllabuses do not coincide entirely pupils may have concentrated on one more than the other, or they may just have been more fortunate with the questions on one paper. There is nothing in this study to throw doubts on the validity of regarding CSE Grade One as equivalent to an 'O' level pass in general. It is in individual cases where the discrepancies occur and may be turned to the pupil's advantage.

However much the place of examinations in schools has been decried, in practice improved educational opportunity has meant greater opportunity to sit, and pass, the examinations which give entry to further education and particular careers. Individual pupils, therefore may find it worthwhile to exploit the two systems and enter for the same subject in both. The following results show what happened to some who did.

Results of double entry

A detailed study was made of seven subject areas. They were the subjects where clear cases of double entry were most common. They are academic subjects. Double entry in art and practical subjects were difficult to identify because of the variations in the subject names, and without a detailed analysis of the syllabuses it would be impossible to determine where overlapping studies were being examined. Table 1 shows what happened when candidates were entered for both examinations.

By adding the passes in the first two columns we can determine the number of 'O' level equivalent passes which might have been missed without double entry. In English, languages, history, and biology these exceed the number of cases where passes at both levels were gained and where pupils may be regarded as examined unnecessarily. It is perhaps easy to understand the large number of double entries in mathematics when its status as a qualifying subject is remembered.

The benefits of backing up the 'O' level by CSE are shown more clearly in Table 2.

While there is no 'near miss' grade awarded at 'O' level, we must assume candidates are better off with a CSE Grade Two than nothing, so for those failing 'O' level and gaining a CSE Grade One or Two, it was a good thing they entered for CSE as well. The proportion of those benefiting varies by subject, but in languages and history it is over half.

Candidates who were entered at 'O' level only and obtained the uncertificated 'near miss' grades should also be taken into account, as they might well have obtained a CSE Grade One or

Table 1: Results of double entry

Subject	*Number of double entries*	*'O' Fail CSE 1* %	*'O' Pass No CSE 1* %	*'O' Pass CSE 1* %	*'O' Fail No CSE 1* %
English (1)	103	19 (18·4)	19 (19·4)	35 (34·0)	30 (29·1)
Maths	233	33 (14·2)	16 (6·7)	111 (47·2)	73 (31·3)
Languages (2)	95	21 (22·1)	5 (5·3)	20 (21·0)	49 (51·6)
History	68	21 (30·9)	5 (7·4)	16 (23·5)	26 (38·2)
Geography	134	9 (6·7)	27 (20·1)	51 (38·1)	47 (35·1)
Biology	109	23 (21·1)	18 (16·5)	35 (32·1)	33 (30·3)
Sciences (3)	194	32 (16·5)	13 (6·7)	62 (32·0)	87 (44·8)

(1) Includes English language at 'O' level and the combined English CSE. The two are treated together as they have similar status as qualifications.

(2) Includes French, German, Spanish.

(3) Includes physics, chemistry, general science.

Table 2: Failure at 'O' level, but success in CSE

Subject	*Number of double entries*	*CSE 1 only*	*CSE 2 only*	*CSE 1 and 2*	%
English	103	19	19	38	36·9
Maths	233	33	40	73	31·3
Languages	95	21	33	54	56·8
History	68	21	19	40	58·8
Geography	134	9	24	33	24·6
Biology	109	23	22	45	41·3
Sciences	194	32	43	75	38·6

Table 3: Numbers taking 'O' level only and obtaining 'near miss' grades

English	52
Mathematics	14
Languages	64
History	46
Geography	35
Biology	17
Sciences	20

Two. The numbers involved are shown in Table 3.

The surprisingly small number of double entries in English is reflected here in the large number of uncertificated 'near miss' while double entry in maths seems to have been a worthwhile precaution. The high failure rate in languages is partly accounted for by the absence of Latin in CSE, and there is perhaps a need for a new examination in Latin.

But the Schools Council advice was that doubtful candidates should be entered for CSE not for 'O' level. If that advice had been followed, the successes shown in Table 4 would have been lost.

Table 4: Success at 'O' level

Subject	*Number of double entries*	*'O' Pass No CSE 1*	*CSE 1 Good 'O'*	*Total Improvement*	*%*
English	103	19	12	31	30·0
Maths	233	16	36	52	22·3
Languages	95	5	4	9	9·5
History	68	5	6	11	16·2
Geography	134	27	32	59	44·0
Biology	109	18	12	30	27·5
Sciences	194	13	19	32	16·5

The value of an 'O' level pass over a low grade CSE is obvious. A 'good' 'O' level is defined as a grade A or C where the Board awards three pass grades (A C E) and four or better where there are six categories. It seems worthwhile for passes at this level to be acknowledged as candidates may well have obtained good 'O' level grades in other subjects and the picture presented by the certificate will be improved by the addition of these successes.

Comparison with Table 2 will show that although the losses from concentrating on CSE would not generally have been as great as relying on 'O' level, there were advantages in taking the 'O' level examination as well. This is particularly true in geography, but applies also in English, mathematics and biology.

To achieve these successes, 936 extra examinations were taken. Shared equally among those who took a mixture of 'O' level and CSE, it is 0·3 of a pass each. In practice it may be as many as three extra passes for one pupil. Even one extra pass may qualify a pupil for entry to a particular occupation or further education course, or avoid the need to resit subjects in the sixth form. By adding together those cases in Table 1 where an 'O' level equivalent pass was achieved in one examination only it is seen that there are possibly 216 extra passes in the subjects concerned. These could have been achieved without double entry if the right decision about which examination system to use had been made, but presumably, these were the difficult borderline cases, or cases where a pass was vital to the pupil's future.

The Schools Council (1971) are hoping for a common system of examination at 16+ by 1977 but while the two systems exist some may benefit from using them both. Whether the extra achievements are worth the extra expenditure of time and money must be for schools to judge. It is hoped that this analysis might help to inform the judgement.

References

Schools Council (1966). *Examining at 16+*. Report of the Joint GCE/CSE Committee. London: HMSO.

Schools Council (1971). *A Common System of Examining at 16+*. Schools Council Examination Bulletin 23. London: Evans Methuen Educational.

King, R. (in press). 'Short-course neighbourhood comprehensive schools—an LEA case study', *Educational Review*.

Question Choice in Examinations

an Experiment in Geography and Science

E. G. Taylor, North Regional Examinations Board
and D. L. Nuttall, formerly Sixteen-plus Examining Project, NFER

Summary

This paper reports an investigation into the difficulties experienced by candidates dealing with examination papers which offer a choice of questions. The study in geography was carried out by asking candidates to continue working the question paper in a different coloured ink after the examination, to see how much of the paper they could answer. In the science study candidates entered for the Board's Mode 3 examinations were asked to work the Board's Mode 1 question papers and to answer as many questions as possible, given virtually unlimited time. The study shows that some candidates may have made a poor choice of questions in geography during the actual examination, and it shows also in this subject, as in science, the extent to which their knowledge was more extensive, if more superficial, than was tested by the restricted choice of questions required. It is significant that, where candidates appear to have experienced difficulties in selecting questions, the greater difficulty should have been encountered among candidates awarded lower grades.

Introduction

It has long been the tradition in British public examinations that candidates be offered a choice of questions in their examination papers. There are two main reasons advocated for upholding this tradition: first, it allows the teacher a certain amount of freedom to select areas of the syllabus most suited to the needs and interests of his pupils and, second, it allows the candidate to show himself in the best possible light by attempting only questions where he is reasonably sure of producing a good answer.

Question choice has its problems; in the words of Christopher (1972):

> Strangely enough it is often thought that the liberal choice of questions usually provided in an essay-type paper ensures that it is a 'fair paper'. In fact, a wide choice of questions ensures that the candidates are being assessed on the basis of what are in effect different examinations. A paper consisting of twelve questions on which any five are to be answered may not appear to allow a great degree of variation but in fact such a paper can produce no fewer than 792 combinations of questions, each differing from the others in at least one respect. (p.5)

Two such papers, which together might form the whole examination, produce 792^2, or 627,264 possible combinations, and combinations in excess of one billion have been noted in current examinations.

If candidates are to have a free choice of questions, then these questions should be of equal difficulty, statistically speaking; that is, if two groups of candidates of equal attainment attempted two different questions, their average marks should be the same. Nuttall and Willmott (1972) in their analysis of some selected GCE 'O' level and CSE examinations demonstrate that this is rarely the case. As a result, it would appear that some candidates can add a few marks to their total simply by being fortunate enough (or wise enough) to attempt the easier questions in the paper, while others are penalized to the extent of several marks by being unfortunate enough to choose the harder questions.

These, then, are some of the problems of measurement posed by choice examination papers, and a fuller discussion of them including an outline of the problems of the validity of a choice paper, may be found in Nuttall and Willmott (1972). As yet, comparatively little is known about them, but even less is known about the problems that confront the candidate when faced with a choice paper. Two studies with university students (Meyer, 1939; Cowan, 1972) have shown that upwards of 25 per cent of students were unable to choose their four best questions from the five that they had answered: they had performed better on the question they chose not to have marked than on one or more of the questions they chose to have marked. This suggests that choosing which questions to answer in the normal situation when the candidate does not have the benefits of hindsight must be at least as difficult a task. It is highly likely that the task will be even more difficult for the average 16-year-old, particularly when he is faced with examination papers that are printed as booklets with up to 40 pages.

The aims of the studies

The two studies reported here aimed to shed light on some of the problems that candidates might encounter in a choice examination. More specifically, the primary aim of the geography study was to discover if candidates could successfully choose, under examination conditions, the questions where they would demonstrate their best performance. The subsidiary aim of the geography study, and the major aim of the science study, was to discover whether a sizeable number of the candidates were able to sustain their level of performance on questions additional to the normal quota, thus demonstrating a wider, though not necessarily a deeper, knowledge of the syllabus than the normal examination would allow. The two studies differed in method and are therefore described and discussed separately. Both studies were, however, similar in that they employed 1970 Mode 1 papers of the North Regional Examinations Board (NREB) and were conducted in Summer 1970 within schools registered with the NREB. Both were also conducted with small and non-representative groups of pupils and under a number of inevitable handicaps, which are discussed below, so that any conclusions can only be tentative.

The geography study

The examination papers

The geography examination consisted of two written papers, together accounting for four-fifths of the final assessment, and a project (or fieldwork or local study), which is not considered further here, worth one-fifth of the total (NREB, 1968a). Paper 1 consisted of a compulsory question on Ordnance Survey (30 marks), a compulsory question on the British Isles map (30 marks), and four questions on regions within the British Isles (each worth 30 marks) from which the candidate had to choose one. Each question, whether compulsory or optional, was of the multi-facet variety, that is, it consisted of up to 30 smaller questions. Each smaller question was compulsory once the candidate had opted for the main question. The question paper itself was 16 pages long and was supplemented by a booklet with 17 pages of maps, sketches and photographs on which the questions were based. The candidates were allowed two hours for the paper.

Paper 2 consisted of one compulsory question on the world map (30 marks), two questions on physical geography (each worth 30 marks) from which the candidate had to choose one, and three questions on natural regions of the world (each worth 30 marks) from which the candidate had to choose one. As in Paper 1, the questions were of the multi-facet variety and the question paper was 16 pages long, supplemented by a 16-page booklet of maps, graphs, sketches and photographs. Two hours were again allowed.

The total mark on the written papers was thus 180, of which half was allocated to compulsory questions and half to optional questions. The apparent complexity of the papers is justified, in the Subject Panel's view, on the grounds of validity: only papers with a large number of short questions can fully and fairly sample the various aims and content of the syllabus in a manner suited to the ability of the average 16-year-old. But there is a danger, which this study sought to explore, that such a format creates different problems for the candidates from the ones they would face with a more traditional extended-answer paper.

Method

A number of schools were asked to allow their candidates taking the examination to be brought back on the afternoon of the examination. These candidates were requested to continue answering the paper (in a different coloured ink) doing as many more questions as they could. This happened on two successive days, for Paper 1 and Paper 2. Candidates were graded on the work done during the actual examination and the additional questions were marked only after grades had been determined.

The obvious drawback of this method was the fact that, although on the first day candidates would come fresh to the afternoon's work, by the second day they would realize what was expected of them.

On neither day were there any controls over the use of reference books in the lunch break. There was no reasonable way round this problem, however, and 12 schools agreed to take part, despite this obvious limitation. Altogether there were 220 candidates from the 12 schools who took part (186 on Paper 1 and 209 on Paper 2).

It is our opinion, endorsed by the balance of members of the Board's advisory Subject Panel, a number of whom had allowed candidates to take part in the study, that the candidates were not as well motivated in the 'experimental' afternoon sessions, particularly on the second day, as in the 'real' morning sessions. Thus any improvements in performance noted in the experimental sessions can be seen as real evidence of difficulties in choosing questions, and the extent of such improvements is likely to be under-estimated.

Results and discussion

Table 1 gives details of the number of candidates showing a deterioration or an improvement in performance from the examination session to the afternoon experimental session, for each of the three optional sections. In all three sections, well over half the candidates did worse in the experimental session, i.e. they showed their best performance in the examination session. A sizeable proportion of the candidates, however, did much better on one question in the experimental session than on the question they had chosen in the real examination. This is most noticeable for Paper 1, which was taken on the first day and which had the widest choice (one out of four questions), over 10 per cent of the candidates achieving at least five marks more on an experimental question than on their real question.

In contrast, for the physical geography questions in Paper 2 where the choice was one out of two, less than two per cent of the candidates demonstrated a similar improvement.

The greatest improvements on an experimental question over a real question were: nine marks (two candidates); eight marks (four candidates); seven marks (nine candidates); six marks (13 candidates).

On the basis of the written papers, provisional grades were awarded under the Board's normal procedures. The CSE grade scale forms the basis of this provisional grading but, in recognition of the many sources of unreliability inherent in any assessment procedure, additional border-zone provisional grades are used at each formal grade boundary (e.g. the border-zone between grade 3 and grade 4 is described as grade 3/4). The border-zones were defined by a mark range of five marks in both directions from the provisional grade boundaries. Thus, if the provisional grade

Table 1: Number of candidates showing various categories of difference between marks achieved on question chosen in examination session and on next best question chosen in experimental session

MARK DIFFERENCE*	PAPER 1 (*Regions of British Isles*)	PAPER 2 (*Physical geography*)	PAPER 2 (*Natural regions*)
+5 or more	25	104	44
0–4 (+^{ve})	85	81	95
1–4 (−^{ve})	55	17	52
−5 or more	21	4	18
TOTAL	186	206	209

*A positive difference indicates better performance in the examination than in the experimental session.

Table 2: Example of mark ranges corresponding to provisional grades

Provisional grade	1	1/2	2	2/3	etc.
Mark range	150–116	115–105	104–96	95–85	etc.

boundary between grades 1 and 2 was a mark of 110, for example, any candidate with a mark between 115 and 105 would receive the provisional grade 1/2. This procedure is illustrated further in Table 2 and gives rise to 11 categories. In general, a change of over five marks in a candidate's total is likely to change his provisional grade.

Thirty-four candidates demonstrated sufficiently better performance (gaining five or more additional marks) on a single experimental question than they did on a real question to cause an improvement in their provisional grade and there were three further candidates with a similar improvement in marks for whom the second paper was not available and hence whose provisional grade could not be determined for the purposes of this study. In contrast, there were seven candidates who showed a similar improvement in total marks but whose provisional grades did not change.

Another group of 11 candidates did not show a large improvement of any one pair of real and experimental questions, but did accumulate an improvement of five or more marks over a number of optional questions in the experimental session. In nine cases, this caused an improvement in provisional grade.

Of the remaining 165 candidates (not all of whom took part in both experimental sessions), 11 improved their provisional grades with an accumulated gain in marks of less than five. Thus some 25 per cent of the candidates taking part in the experimental sessions were able to improve their performance sufficiently to change their provisional grades. Of course, the percentage whose *final* grade would be changed is much smaller since there are only six official grades, and other evidence in the form of a project or fieldwork is also taken into consideration. (For example, a candidate who moved up from provisional grade 4 to provisional grade 3/4 might well still have received a final grade 4.)

Table 3 gives a breakdown of the improvements in marks obtained by candidates in each provisional grade. It is apparent that more candidates in grades 3 and below improve their total by at least five marks than do candidates in higher grades.

This study clearly demonstrates that some candidates in the sample failed to choose the optional question on which they were capable of producing their best answer in the actual examination situation. A substantial proportion showed themselves capable of producing almost as good answers on questions which they had not chosen (and had no need to choose) as on the questions they chose in the examination proper. But it must not be forgotten that the majority of candidates did better in the real examination than in the experimental session, showing themselves capable of making a wise choice of questions.

The science study

The science study investigated candidate's performance in physics, biology and general science. The same experimental design was used for each subject, and since the results were very similar for the three subjects, only the physics study is reported here in any detail.

The examination papers

All the major science subjects have the same basic format: the assessment for each subject consists of a common core science paper (the same paper is used for all the subjects), accounting for

Table 3: Numbers of candidates improving their marks by provisional grade

IMPROVEMENT IN MARKS	GRADE 1	1/2	2	2/3	3	3/4	4	4/5	5	5/U	U	TOTAL
10 or more	–	–	1	1	2	3	2	1	–	–	–	10
5–9	–	1	2	6	5	14	7	3	2	4	1	45
0–4	7	18	18	31	24	26	18	13	3	5	2	165
Total	7	19	21	38	31	43	27	17	5	9	3	220

one-fifth of the total, an assessment by the teacher of course work and practical work also accounting for one-fifth of the total and a specialist subject paper which contributes the remaining three-fifths to the total. Candidates also have the option of submitting a project which can serve to enhance their final grade by a single grade, but never to reduce it (NREB, 1968b).

This study employed only the specialist subject paper. In the case of physics, this consisted of a 20-page question booklet in which the candidates had to record his answers, containing 10 questions of which the candidate had to attempt any five. (This paper thus contained 252 possible combinations of questions.) As in geography, each of the main questions contained about 10 or more sub-questions; once the candidate opted to do a particular question, he was obliged to attempt all the sub-questions: in other words, there was no choice element at the second level. Each question was marked out of 40, according to a detailed mark scheme, and two-and-a-quarter hours were allowed.

Method

The method adopted in the science study was markedly different from that of the geography study. Instead of requiring candidates to continue answering their own papers in an experimental session, candidates for physics under Mode 3 were asked to sit the Mode 1 physics paper, and to answer as many questions or parts of questions as they were able. The Mode 1 paper was set on a different syllabus from the one followed by the candidates and the type of question paper was not necessarily familiar to them; in physics, however, it is likely that there would be considerable overlap between the aims and content of the Mode 1 syllabus and the aims and content of the majority of school-based Mode 3 syllabuses. One or two schools which had been invited to take part did, however, decline to do so as they felt that their Mode 3 syllabus and approach to the subject differed substantially from those of Mode 1 and hence that the Board's Mode 1 question paper did not offer an appropriate means of examining their syllabus. One school which did take part commented later on the unsuitability of the question paper for the course their candidates had followed. The major problem was, however, the lack of motivation as it would have been contrary to the Board's policy to allow the Mode 1 question paper to contribute to grades in Mode 3.

Fifteen schools, with a total of 280 candidates, took part in the physics study. Up to three hours were allowed (the normal time being two and a quarter hours) but in the event few candidates took more than two hours to complete all they could. This provides further evidence both of lack of motivation and of the unsuitability of the question paper. In biology, 15 schools with 228 candidates took part and in general science, three schools with 41 candidates.

The questions were all marked, and the marks on each question were arranged from highest to lowest. Provisional grades, similar to those in the geography study, were determined on the basis of the candidates' performance on their best five questions.

Results and discussion

Table 4 gives the breakdown of mark differences

Table 4: Number and percentage of candidates demonstrating various differences in marks between their fourth and fifth, fifth and sixth, and sixth and seventh best questions

MARK DIFFERENCE	FOURTH AND FIFTH		FIFTH AND SIXTH		SIXTH AND SEVENTH	
	Number	*Per cent*	*Number*	*Per cent*	*Number*	*Per cent*
5 or more	21	7	26	9	20	7
4	17	6	20	7	16	6
3	37	13	26	9	16	6
2	58	21	54	20	59	22
1	81	30	79	29	91	33
0	64	23	72	26	70	26
*Totals**	278	100	277	100	272	100

*The slight difference between the totals reflects the fact that not all the 280 candidates attempted five questions; of those that did five questions, not all attempted a sixth, and so on.

Table 5: Number of candidates demonstrating various differences in marks between their fifth and sixth best questions, by provisional grade

MARK DIFFERENCE	PROVISIONAL GRADE									TOTAL
	1	1/2	2	2/3	3	3/4	4	4/5	5 and below	
5 or more	2	1	3	4	6	3	6	1	–	26
4	1	–	1	1	1	8	4	2	2	20
3	1	1	–	2	3	3	10	2	4	26
2	3	2	2	7	8	11	7	8	6	54
1	4	2	6	8	4	8	32	11	4	79
0	3	2	3	5	8	10	19	7	15	72
Total	14	8	15	27	30	43	78	31	31	277

obtained between the candidates' fourth and fifth best questions, their fifth and sixth best questions, and their sixth and seventh best questions. For example, nine per cent of the candidates obtained at least five marks more on their fifth best question (which counted towards their provisional grade) than on their sixth best question (which did not); in contrast, 26 per cent of the candidates did equally well on both and a further 29 per cent obtained only one mark more on their fifth best question than on their sixth best question.

It is clear, therefore, that a substantial proportion of the candidates were capable of doing equally or nearly as well on questions which they would not have had to choose in the real examination situation as they did on the fourth and fifth questions they would have had to have chosen. This suggests that some candidates might have found it difficult to make up their minds about which questions to choose; in a sense, they would have been spoilt for choice. It was rare, however, for candidates, especially those of lower ability, to complete any question.

Unlike the case of the geography study, the choice problem appears to be as acute for candidates in the higher grades as it is for candidates in the lower grades. Very similar pictures emerge for the comparisons of fourth and fifth best questions, and of sixth and seventh best questions, which are hence not shown here.

A comparison was made between a candidate's grand total of marks for all the questions attempted and the total mark on his best five questions. The correlation coefficient (Pearson product-moment) between these two marks over all candidates was 0·97, and in biology (where the comparison was between the grand total mark and the mark on the candidate's six best questions) the correlation was 0·96. These correlations, which are artificially boosted in that the mark on the best five questions is also a major component of the other variable, the grand total mark, are nevertheless high enough to suggest that, *if* a candidate is able to select his best five questions, the same rank order would be obtained whether the examination was given under normal untimed conditions with choice limited to five questions or under untimed conditions where the candidates were asked to attempt all questions. The geography study suggested, however, that candidates were not always capable of selecting their best questions.

The correlation coefficient conceals differences between candidates at different ability levels; whereas among candidates in grade 2 or above there is a noticeable decline between the average mark on the best five questions and the average mark on the remainder of the questions, this is not the case for the candidates in lower grades. This again suggests that almost all questions could be attempted at the same level of competence by candidates below grade 2.

Conclusions and implications

Any examination can only hope to sample the knowledge and skills that the candidates should have acquired during the course. To be valid, the examination should be comprehensive and systematic in the way it samples the knowledge and skills of the candidates, but no examiner would ever claim that his examination provided a complete picture of all that the candidates had learned during their course.

In one sense, the results of both the geography and the science studies are encouraging: they indicate that the majority of candidates are capable of doing more than the examination requires them to do, at very much the same level of competence. In other words, the examinations appear to be assessing a fair and representative sample of the candidate's knowledge and skills. But, for choice examinations of the type studied here, this same finding is disquieting in another sense: a sizeable minority are likely to find it difficult to choose the questions where they can show themselves to best advantage. The geography study showed that candidates were often able to do rather better on questions which they had not chosen. Although the differences in terms of marks were often slight they were sufficient to cause an improvement in provisional grade for about 25 per cent of candidates.

It proved to be very difficult to plan the studies and in neither study do the sampling procedures allow of any tests of statistical significance. The results refer, in the first place, only to the scripts produced by candidates taking part in the two studies. Nevertheless, the results are sufficient to suggest that choosing questions in papers of this type presents problems for candidates over the whole CSE range of ability, and particularly for candidates in the middle of the range and towards the lower end. It is at this level of ability where a considerable increase in the number of candidates is soon to be expected as a result of the raising of the school leaving age.

It is important to note that the problems have been studied here in subjects in which those responsible for the CSE in the NREB have departed substantially from more traditional types of question papers in an attempt to achieve greater validity in the examination. There is no claim that candidates taking a traditional essay-type paper face these problems of choosing questions to anything like the same extent. It would be worth mounting a similar study with such papers to explore the problem further.

Of course, the problem of question choice discussed here is only one of the sources of potential misgrading. Earlier studies in marker reliability (NREB, 1970) indicated the need for wide border-zones at the provisional grading stage; in practice the unreliability of the examination is mitigated, to some extent, by the Board's policy of including elements other than written papers, in the scheme of examination. The unreliability of marking complicates the interpretation of the results of the present studies. Had the scripts been marked by a different marker, or by the same marker on a different occasion, some of the marks would have been different. This implies that some candidates who were found to have performed slightly better on a question in the experimental session than on a question in the real examination might not show the same difference on another marking occasion. Over all candidates, the net effect should, however, be negligible. Nevertheless, in the wider context, the fact that a candidate's mark on a question is dependent both on the marking scheme and the marker means that it is unlikely that every candidate could always choose his best questions even when he has written answers to all questions, as the study by Cowan (1972) suggests. But how can the problems of question choice be eliminated or reduced?

In the science subjects, the instructions to candidates might be modified as follows:

(a) You are expected to be able to answer five questions in the time allowed.

(b) Attempt the five questions which you can do best in the time allowed; try to *complete* these questions.

(c) If you answer more than five questions, you will receive credit only for your best five answers.

A more fundamental change would be to instruct candidates to answer as many of the questions as they felt they were able to attempt, even if they did not complete any or all of them. The most fundamental change would be to make the questions shorter and all compulsory. This has obvious advantages for the reliability and comparability of measurement, since one would no longer have hundreds of miniature examinations within the one paper, but its educational advantages are not so apparent; many teachers would feel that their freedom to concentrate on specific parts of the syllabus according to the needs and interest of their pupils was being inhibited.

In geography, the structure of the papers is more complex, designed to ensure that candidates answer questions on Ordnance Survey, British Isles, physical geography and world natural regions. Again, an interpretation of the instructions similar to that suggested for science would mitigate the problem of appraisal and choice. The use of shorter questions, covering a wider variety of topics and supported by the instructions to 'answer all questions' might also be considered and, since the element of choice within the syllabus is considerably less than in science, might be more acceptable to teachers. Discussion of the results of this study has led to proposals within the examining board concerned for alterations to the structure of question papers and to instructions to candidates, in a number of subjects.

References

CHRISTOPHER, R. (1972). 'The prospect for examinations', *Dialogue* (Schools Council Newsletter), 10, 3–5.

COWAN, J. (1972). 'Is freedom of choice in examinations such an advantage?', *The Technical Journal*, February, 31.

MEYER, G. (1939). 'The choice of questions in essay examinations', *J. Educ. Psychol.*, **30**, 161–71.

NORTH REGIONAL EXAMINATION BOARD (1968a). *Syllabus for the 1970 Geography Examination.* Newcastle-upon-Tyne: NREB.

NORTH REGIONAL EXAMINATION BOARD (1968b). *Syllabus for the 1970 Physics Examination.* Newcastle-upon-Tyne: NREB.

NORTH REGIONAL EXAMINATION BOARD (1970). *Marker-Reliability: Geography and Physics-with-Chemistry.* Notes for Teachers, No. 10. Newcastle-upon-Tyne: NREB.

NUTTALL, D. L. and WILLMOTT, A. S. (1972). *British Examinations: Techniques of Analysis.* Slough: NFER.

Education 3-13

A new journal for Primary Education

In the April issue

Headteachers
Do we need Headteachers?
Aims, organisation and planning: an account of practice in one Middle School
Talking to Headteachers
A School re-organises
Plowden aims: a review

It may be ordered direct from the Publishers

In Britain

The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada

Agathon Publications Services Inc.
150 Fifth Avenue
New York, N.Y. 10011

Annual subscription £1.00 including postage

Book Reviews

Classrooms Observed

by Roy Nash
Routledge and Kegan Paul, 1973, 138 pp.,
Hardback £2·50, Paperback £1·25

In *Classrooms Observed* Roy Nash provides a valuable study of classroom interactions, in an attempt to investigate the factors affecting children's learning. It is both complementary and supplementary to large scale survey research in that it looks at children's learning through a microscope rather than a telescope. Dr. Nash spent three years in the systematic observation of teachers and pupils, mainly in the lower forms of a Scottish comprehensive school, and in the five junior schools which supplied its pupils. Not only did he watch what they did, but he obtained further information from discussions with the teachers and pupils, and from tests, questionnaires and school records.

The classroom 'ecology' is considered in relation to the set of personal constructs held by each teacher. These constructs were uncovered by discussing with the teachers how they differentiated between individuals in their classes. Different teachers, naturally enough, had different sets of constructs. Three which were common to eight primary school teachers were the following: *hardworking-lazy; mature-immature;* and, *well behaved-poorly behaved.* In contrast, secondary school teachers shared only *bright-dull; lively-lumpish* and, *likeable-less likeable.*

Several interesting findings emerge, some confirm teachers' intuitive knowledge about education, others shatter a few beliefs.

In the schools studied, the primary schools tended to devote more time to 'written seatwork' than did the first year of the comprehensive schools. Dr. Nash suggests that this might be due to primary teachers over-preparing their charges for what they (mistakenly) think is the secondary school style of teaching, and conversely, secondary teachers accepting a less formal style because they think that this is what new pupils are used to. In fact, in this study it was the pupils from the most formal primary school who had most difficulty in adjusting.

The study brings out in detail the way in which teachers' perceptions and expectations affect the pupils' achievements. This has been noted in several large scale surveys, but Dr. Nash tried to isolate the ways in which this was brought about. Teachers tended to see pupils in terms of personality attributes, rather than ability. Thus, less able children who were seen unfavourably tended to be transferred to remedial groups rather more often than favourably regarded children of equal ability. Regardless of the class or school organization, children were remarkably accurate in estimating their relative positions (and those of their classmates) in the academic hierarchy as perceived by their teachers.

Dr. Nash also demonstrated that in this sample, social class was of less importance than was the children's friendship patterns. Cliques within the classroom shaped the children's self-image, which, in turn, moulded their performance.

He argues that if socially deprived children are unable to communicate because of language barriers it is teachers who should be able to adjust their own use of language more easily in order to teach.

This is a refreshing book. It should be found on each staffroom table, and be compulsory reading for all those about to teach or conduct research into teaching.

We have watched the school island through our telescopes from the ship anchored in the bay long enough. Now that one of the explorers has returned with the news that the natives are friendly (even if slightly misguided), let us have a few more landing parties. There are many more native villages to explore.

ROLAND HOSTE

Learning to Teach

by Brenda Thompson
Sidgwick and Jackson, 1973, 208 pp., £1·95

Brenda Thompson, herself an experienced teacher and now a headmistress, has written her new book *Learning to Teach* primarily for students and probationary teachers. It is designed as 'a practical guide to the day-to-day running of a class in a primary school'. She does not intend it to be the complete guide to teaching but a guide to help new teachers with the pressing problems of organization, control and discipline.

Although important issues like what is taught in schools are not dealt with, at least Mrs. Thompson hopes that by using the advice, help and practical tips which abound in the book, new teachers will be able to organize their classes competently enough to be able to teach something, a situation which will not occur if she is unable to control the children sufficiently to make them listen. Mrs. Thompson's concern, then, is to help her audience to reach professional competence in this area of their work.

That they must reach this level of competency quickly nowadays is necessary because, as the Plowden Report points out, 'a revolutionary change has come over the composition of the body of teachers'. Whereas before the war 'a school staff was a body of experienced professionals in which a newcomer could easily learn to find his feet', today 'a small body of experienced teachers is surrounded by a rapidly changing group of young women'. Mrs. Thompson believes that experienced teachers are unable to articulate and pass on practical advice and so, if young students and inexperienced teachers are unlikely to get advice from their experienced colleagues, where can they turn to get help?

One might suppose that Colleges of Education would be concerned with preparing students for the kind of real life classroom situations for which Mrs. Thompson is offering her sound advice, but it is generally known and accepted that most colleges of education do not give much practical assistance to their students. Reasons are offered in the book as to why students leave college ill equipped to teach, one being that the experiences of the college staff themselves, limited perhaps to either the old or the new teaching methods in schools with a particular type of organization, means that they are personally acquainted with only a small sector of the possible range of different forms of organization found in primary schools today.

The author is very aware of the nature of teaching, saying that 'no other job is so harshly demanding, requiring the recruit to turn overnight into an expert', and she sees her book as an aid in this process, giving the tricks of the trade, and designed to ensure riot-free classrooms for the followers.

Perhaps it is too much to hope that students will in future get this kind of very practical, sensible advice, like techniques to use if you loose control of the class, kinds of stories to read and how to read them to a new class, from either their college course or from teachers supervising them on teaching practice or in their first appointment. Until they do, or if they ever do, then I think that Mrs. Thompson's book fills the gap admirably.

BARBARA Y. ROBB

Education and the Severely Handicapped Child: Subnormality in the Seventies

by Peter Cummings

National Society for Mentally Handicapped Children, 1973, 68 pp., £1

During the past twenty years the field of mental retardation has become a complex quarry. Many disciples are exploring it. Whereas a generation ago a solitary student could have reviewed all the relevant literature by himself, today he would have difficulty in keeping up with the relevant literature from his own profession.

This 60 page monograph is accordingly limited in aim, but is compact, instructive and challenging. The first twenty pages 'review literature pertaining to the education and training of severely mentally handicapped children'. In this context 'training' and 'education' are seen to be distinguished by staff for whom the first is a ward activity and the second concern of the teacher. For a further twenty or so pages data provided by questionnaires are analysed. This no doubt illustrates practice in far more than the twenty-five subnormality hospitals surveyed. At first sight the discussion section which follows, may suggest that the author has chosen an easy exercise of doubtful value. To compare *practice* in subnormality hospitals with the more positive findings of research on learning, teaching methods, speech and language, motivational factors, and so on, leads on to easily predictable conclusions. But realism and concern and not cynicism prompt the author to ask, 'who writes for whom?' or, 'who reads it when it is written?'.

What may appear to be a series of depressing conclusions point instead in the direction of required effort. As the author himself is aware, the date of the survey of the subnormality hospitals could increasingly prove relevant; it was one year before the 1970 Education (Handicapped Children) Act was passed and two years before it became operational.

The 1969 figures offered in this study and the 'pertinent comments by staff' now offer us the alternatives of 'a remainder of the doubtful old days or a standard to aim for'.

This book is particularly commended to those who help train staff for subnormality hospitals.

S. S. SEGAL

Selection for University in Scotland

by John L. Powell

University of London Press, for the Scottish Council for Research in Education, 1973, 104 pp., 80p

An account is given in this book of the Assessment for Higher Education project, undertaken on behalf of the Scottish Council for Research in Education. Although the project was concerned with selection-procedures affecting the whole field of higher education in Scotland, thus including the Scottish central institutions and the Colleges of Education, this volume concentrates on problems affecting the university area.

The research is welcome for it throws further light on the predictive efficiency of measures which may be used to identify potential student success at university; a comparison is made of the forecasting value of Scholastic Aptitude test results, with performance in the Scottish Certificate of Education examination (SCE), and with teacher-estimates of candidate's likely achievements at university, the student-wastage problems was one of the primary concerns of the study. Another consideration appears to have been the possibility of freeing the secondary school curriculum from the restrictions, said to be imposed on it by the SCE—provided that a satisfactory substitute could be found for this cognitive measure.

The methodological approach is thorough and, in some respects, comprehensive. The sample includes nearly all pupils in Scottish schools presented for the SCE (Higher Grade) for the first time in 1962. This involved data-collection and processing on 11,600 pupils from 218 schools. Of these, the progress of 2,781 students entering universities was recorded on a longitudinal basis until their courses terminated. The categorization of the predictor-variables is equally detailed in both quantitative and qualitative terms, and the statistical techniques employed, as well as the findings, are reported with clarity.

The results should be of no surprise to anyone familiar with these issues. The Scholastic Aptitude Tests proved to be of little value in predicting the achievement of a degree, nor were they of much value as supplementary measures. Neither teachers estimates of performance in individual subjects in the SCE nor headteachers' estimates of success at university appeared to be as good predictors as performance in the examination itself—though it is pointed out that these findings should be treated with caution. Although the SCE turned out to be the best pre-entrance academic predictor, its value in forecasting success in university terms is regarded as less than satisfactory by the author. Performance in first-year university examinations, by contrast, turned out to be a much more reliable guide to the future.

It will have been noted that the study is concerned with cognitive measures—mainly tests of ability and tests of attainment. Perhaps this limits the value of the research but it was based on a deliberate policy-decision. Factors such as personality, social background, health, work-habits, financial security and so forth, which may be of some significance in the successful completion of a degree course, were not considered not only because such data are difficult to quantify and handle in an objective manner, but also because of what the author regards as the 'public unacceptability' of these as criteria for selection. (A curious reason!) In this connection it is interesting to find that the book ends on the note that '. . . effort should be directed to seeking good non-academic predictors if only because the chances of improving academic ones seem so poor' (p. 81). Perhaps, therefore, future investigations should be more wide-ranging and take into account research findings on attitudinal aspects and related matters. One final point. It is a pity that the study has a parochial flavour. After all, Scotland is uniquely placed since a significant proportion of the student population has been exposed to different pedagogic and examination patterns in the final stages of secondary schools to the ones considered in this investigation. One wonders, therefore, how, the GCE 'A' level would compare with the SCE as a predictor of success at university, whether the forecast ratings of English teachers are closer to the mark than those of their Scottish colleagues, whether differences in curricular content and method may not be of some relevance to issues of selection. It is just possible that the author's view regarding the difficulty of improving academic predictors is overpessimistic.

FRED HONIGMANN

The Coming Crisis of Western Sociology

by Alvin W. Gouldner

Heinemann Educational Books, 1971, 528 pp.
Paperback £1·50

'The Coming Crisis' was first published in 1971 and was recognized at once as a major sociological work. This review is therefore, in effect, a retrospective assessment of its virtues and impact. Gouldner's book expresses the academic and intellectual ferment that every student of sociology knows is going on within his discipline. But it would be unfair to Gouldner to say that he was only voicing what everybody knew already. His book brought the sociological ferment together in a major work and at an early date. As far as this review is concerned for 'sociology' one could read 'social science' or 'educational research', for Gouldner is, in essence, questioning assumptions that lie behind the conduct of empirical research as much as the construction of theories.

The book itself is very readable. Gouldner's style is lightweight and persuasive, though the content at times is heavy stuff. It is, however, an unbalanced work and the reader must wait until page 341 before the kernel of Gouldner's argument is fully elaborated. Before this the reader is given a thorough tour of 'The World of Talcott Parsons' (pp. 167-372). Gouldner writes of 'at least one important sociologist, a man of understanding and enduring attainment, who was so traumatized by his difficulty in understanding Parsons' work that he assumed this signified his own impending obsolescence' (p. 201). While not of the same status as 'one important sociologist', I, and many others, know the feeling of depression and inadequacy that comes from attempting to read Parsons' voluminous works. Gouldner is therefore recommended to these students who feel they ought to read Parsons, because he is said to be the major sociologist of the Western world, but obtain little from the exercise except a dull ache in the brain.

The emphasis upon Parsons, however, means that other writers who have attempted to break from functional explanations, Goffman and Garfinkel for example, barely get a mention. And surely Marxist sociology, both in the Western world and as the academic sociology of the Eastern bloc, deserves more elaboration?

The basis of Gouldners' thesis is that everyone makes assumptions about the natural and social world, the way it is as fact and the way facts interrelate. Such assumptions provide our orientation to the world and define it for us. Of course there must be a congruence between our background assumptions, our personal reality, and the world. An absence of congruence will produce a cognitive dissonance. What Gouldner points out is that social scientists have background assumptions too, for they are also human. At a very generalized level assumptions are a guide to everyday living; at a more specific level they enter into the domain of each particular activity, and, of course, social science is one such activity. Hence our assumptions about the world are unconsciously built into our social theory. Academic sociology, the sociology of the universities of the Western world dominated by Parsons' influence, has the built-in assumptions of a middle class utilitarian culture. Gouldner argues that Parsons was able to maintain a congruence between his sociological theory and the way he perceived the world even through pre-war economic depression, as he (Parsons) was safely wrapped up in the protective cacoon of a high status university. However, today things are different and the 'real' world has caught up with the ivory-towered theorist. There are three reasons for this implosion upon the theorist. Firstly, the students, the middle class drop-outs and adherents to the psychedelic philosophy whose values are in contradiction to the utilitarian culture. Secondly, the mammoth financing of social science by the state to solve problems from Watts to Vietnam. Thirdly, the rise of the welfare state. These last two reasons, according to Gouldner, challenge the assumption of functionalist theory that society transcends individual men and automatically stabilizes around a dynamic equilibrium. The deliberate channelling of finance to social scientists for the solution of contemporary problems, and to the population at large through the welfare state, interferes with this 'automatic' stabilization, thus contravening one of the basic tenets of the functionalists' theory. This challenge has therefore changed the assumptions and consequently the theory of academic sociology. For Gouldner this accommodation of academic sociology to the implosion of the 'real' world has manifested itself by the embrace of elements of Marxist sociology, particularly in relation to theories of social change and the role of the state. Similarly, with the 'liberalizing' of the soviet regime and the increase in production for consumption, the Marxist academic sociology there shows signs of embracing elements of functional sociology. Gouldner therefore sees a convergence between Western and Eastern sociology, although a distinctive new sociology has yet to emerge.

I am impressed with Gouldner's thesis but there are two points on which I disagree. Firstly, Parsons' theory would appear to be able to account adequately in its own terms for the phenomenon of state interference in the workings of society. This would simply be seen as the state becoming part of integrative and adaptive institutions of sociology. Secondly, there is a danger in assuming that the adoption of apparently Marxist elements of explanations by academic sociology, and of functional elements by Soviet sociology, signify any radical change in the background assumptions of social theorists. As Mezaros notes, Rousseau's analysis of the alienation of man from nature is along Marxist lines, but the analysis is incomplete and the solution is definitely non-Marxist. Rousseau was no Marxist; no more are the academic sociologists of the West. There are elements of Marx but there is not necessarily a drift away from their utilitarian values. The reverse holds true for Eastern sociologists.

What Gouldner is offering is a convergence of social theories, similar to Smelser and Moore's convergence theory of social change, i.e. all industrial societies become the same as they have to obey the imperatives that shape the form and content of industrialized society. However, the convergence theory ignores the ideology of society (the collective assumptions) and assumes that change can be only in one direction. The weakness is in the assumption that theory will necessarily change in a particular direction.

G. Hudson

The School Counsellor

by Ken Williams

Methuen, 1973, 126 pp., Hardback, £2. 00,
Paperback 95p.

When, as part of a recent research project, a head was asked his views on having a school counsellor, he replied: 'I don't want anybody like that here meddling about with everybody'. He, and others who share his attitude, may find little to allay their fears in Ken Williams' recent book, for 'meddle' a counsellor must. Indeed Mr. Williams cites a sad example of one counsellor who, through timidity, didn't, and found as a result that he had no clients. For practising or intending counsellors, however, together with teachers working with them, and those thinking of introducing a trained guidance specialist to their schools, this book should prove of interest.

The author, himself a teacher and counsellor, covers a wide field. The book starts with an examination of the counselling role and proceeds to demonstrate the need for counselling in a rapidly-changing social structure. Succeeding chapters deal with the nature of the counselling relationship and its application to a variety of group and individual situations, to the teacher in his classroom and to the children themselves in their counselling of one another. It is perhaps indicative of an as yet insecure profession that the claims made for the counsellor's skills and training at times appear laboured. The author catalogues these at some length and in the first chapter we are told that, as well as covering the three main areas of vocational, educational and personal counselling, all counselling courses have core studies in the sociology and psychology of education and provide 'extensive' practical experience. The counsellor is also trained in the use of objective tests and has 'expertise' in scientifically-constructed measurement techniques. It should perhaps be pointed out that all such skills are generally acquired in one year. The breadth of the counsellors' scope is certainly well-reflected in the variety of sources from which this book culls support; Plato, Juvenal, Locke, Rogers, Bernstein, Peters, Laing, Hoggart—and many others—are all there. It is perhaps inevitable with such an array that the synthesis is occasionally superficial, as for example in the discussion of the behaviourist criticisms of the Rogerian counselling model.

The chapters are enriched by a wealth of illustration and example taken from the writer's experiences in this country and in the States. There are indeed some sad stories and these are generally presented without a surfeit of sentimentality. In some instances, however, the cynical reader may have difficulty in suspending his disbelief, as when the author records his experience of the group work of Ohlsen in Indiana. A problem faced by one of the groups was that of persuading a black male, who had recently come to the group, that he was accepted:

> Ohlsen explained to the group that the words would never solve this problem. He asked, 'How can you make this guy feel that he is

wanted and not looked upon as a load of crap?' One of the girls suggested giving him a kiss. 'Great, said Ohlsen, 'Let's all cuddle him', and so they did, not once but every time he came to the group. Everyone, female and male including Ohlsen swamped that boy with hugs and kisses and his bitterness melted. He found it impossible to sustain hatred towards people who so demonstratively loved him.

There is much of practical value in this book and the author presents a volume of material which may prove useful to the counsellor confronted with day-to-day problems in school and with the fact that he frequently has to work with a very large number of children whose confidence must be won before a counselling relationship can develop. Mr. Williams is sensitive, too, to the counsellor's position in relation to his teaching colleagues. The point is made that what a counsellor can achieve may be largely determined by their understanding and acceptance of his role; it would have been interesting, however, to have had more examples of how such a role can be modified by their expectations in different school circumstances. The reader forms the impression that the author's primary concern is with personal counselling, and the educational and vocational aspects of school counselling receive sparse attention, although evidence is given that it is in these areas that a school may wish its counsellors mainly to operate. Mr. Williams quotes, for example, that of nine jobs advertised in one month in the *Times Educational Supplement*, only two were concerned with personal counselling whilst three were for work in educational guidance and four were to be exclusively concerned with careers advice. Despite the neglect of these two areas however, the book provides an extensive overview of many of the major problems and issues in counselling, as well as an indication of its theoretical background. It is well-written and illuminated throughout by the author's compassion and commitment.

MARGARET REID

Conflict and Change in a Technical College

by Beryl F. A. Tipton

Hutchinson Educational, 1973, 131 pp., Hardback £2·25

Over the last decade, major changes have taken place in the field of further education and the danger always exists of research becoming obsolete on completion. In this book, the placement of technical colleges as having 'almost complete responsibility for post-school formal education and training for industry, commerce and allied occupations' is perhaps not as meaningful currently as five, or even ten years, ago, particularly when considering degree and HND-level work in the context of the rise to prominence of polytechnics. This, however does not detract from the value of this book which seems to lie principally in a clear and detailed analysis of the evolution of a present-day technical college in terms of its organization and staff structure.

Careful delineations are made of the various roles played, particularly by members of staff, in determining the velocity of growth of the technical college; that is to say its rate of growth, and the directions in which growth occurred. The conclusions must offer some cause for concern particularly to administrators, policy-makers and the most neglected object of research in further education: the student. The author has wisely declined to comment on the benefit, or otherwise, accrueing to students, from the consequences of the 'competition' and 'rivalry' between individual members of staff and departments which is apparently such a striking feature of the climate within the college.

Certainly the most disturbing point to be raised by this research is the fact that: 'it (the college) was not averse to running down its courses even when they were enthusiastically supported by the market, or energetically pushing courses for which demands were dubious or to which other institutions were laying claim'. It is clearly shown 'that the reasons for this can be found mainly in the calculative' outlook displayed by the majority of the staff, which was combined with a general tendency for the extrinsic aspects of teaching to be regarded as more motivating than were the intrinsic ones, and in the wide scope afforded to college development by initial government directives.

Certain problems seem to have existed with regard to the interviewing schedule despite the fact that the questions were, in most cases, merely 'guidelines' for the interaction that occurred. For example, responses to some questions seem to have been omitted completely and the construction of some of the more open-ended questions leaves some room for doubts as to their purpose and worth in this context: 'The questionnaire included an open-ended question asking . . . , but in the end we decided against using these answers for a conventional statistical table (because; footnote). It was found that respondents needed a great deal of help with prompts'.

Overall, the author has shown a face of the technical college that seldom comes to the notice of those concerned with education, and with educational research in particular. In general, the use of documented opinions and interview responses presents a refreshing insight into what is obviously a very complex organizational structure. As indicated in the summary, 'it is hoped that more studies of colleges will be undertaken' and perhaps one could hope that these will not be restricted to technical colleges.

S. KURLE

School Organisation and Pupil Involvement

by Ronald King

Routledge and Kegan Paul, 1973, 256 pp.
Hardback, £3·50

This book is an important one, firstly, because of its methodology, which goes beyond the description of the school as a bureaucracy and incorporates empirically derived measures of organization, and secondly, because of its findings.

It sets out to answer three questions:

(1) How do schools organize the behaviour and learning of pupils and how may they be explained?

(2) To what extent and in what ways are pupils involved in their school?

(3) To what extent or in which ways is pupil involvement related to the internal organization of the school?

In relation to the first of these, data were collected from a representative sample of seventy-two secondary schools from the south-west of England. Over one thousand items of organizational data were collected from each of the schools under three main headings—instrumental and expressive activities and performance, internal and external contextual factors and various structural considerations. In addition to numerical and percentage data, some of these organizational items were used to construct fifty-four unidimensional scales. These scales were based on the work of Levy and Pugh in industrial organization who initiated the use of item analysis in the construction of organizational measures. Unfortunately many of the scales are short—the number of items range from two to eight. King gives as the reason for this, the problem of identifying sufficient numbers of conceptually similar items which discriminate amongst schools. It would have been interesting to see to what extent the conceptual constraints he imposed upon himself were justified in empirical terms at the level of the items. Nevertheless this methodology is undoubtedly a useful one which is probably capable of development and greater sophistication.

The interrelationships amongst the organizational variables were examined under a number of headings in various chapters of the book—the expressive order, school rituals, the house and pastoral care system, games, the instrumental order, ability specialization, age specialization, authority relationships and sex specialization. It is not possible to summarize the multiplicity of findings here, but they are of interest for they provide empirical evidence about aspects of English schools which at the present time, tend to be treated at the level of intuitive beliefs or as inferences derived from American or industrial studies. Some of these beliefs are confirmed e.g. 'attempts to democratize schools seem to be limited exercises' p. 154, and some are not confirmed e.g. 'the increase in the number of rules and set procedures, and the increased use of paper are clearly not functional requisites of large schools' p. 167.

King goes on to offer an explanation of the modes of organization identified in schools. This explanation is presented in terms of the ideologies of teachers and head teachers which he has derived from interviews with heads, analysis of school documents, his organizational analysis and his own experience. This 'explanation' is not a major part of the study, the treatment being limited to nine pages of the text.

The second question, about pupil involvement he examines through a self-report questionnaire which asks for data about extra-curricular activities and responsibility, and presents a number of attitudinal items to a random sample of over seven thousand pupils drawn from a sub-sample of thirty schools. One general and six specific Thurstone-type scales are derived to quantify the pupils' affective evaluation of the school. Although the general scale is sufficiently reliable for use with group data, the specific scales are perhaps less adequate in this respect, four of them having fewer than three items. In some cases, however, other self-report data compensate for this limitation.

Within schools, differences in the degree of involvement are shown related to the social characteristics of pupils, but *between* schools neither internal nor external contextual variables are directly related to the pupils' general level of involvement. The same conclusion emerges with regard to organization: with the exception of streaming in some schools, variations in organization are not related to pupil involvement. This empirical finding, given the advance in methodology which the book represents, thus makes a significant contribution to the debate about the role of organizational factors in relation to the realization of school goals.

Hence the book should be useful to educationists who are interested in coming to grips with the complexity of schools and who have grown weary of the apparently endless ideological claims and counter-claims propagated by the proponents of one or other mode of organization. Because of its methodology, the book will also be useful to researchers in colleges and universities concerned with studies of the school as an organization.

DOUGLAS FINLAYSON

Journal of Curriculum Studies **Collins**

In the May issue

Ideology and the Curriculum: the Value Assumptions of System Builders. *Fred Inglis*

The Primary School Curriculum in a Colonial Society. *M. Kazim Bacus*

Some Problems Raised by the Schools Council Humanities Project. *Sally Wringe*

Theoretical, Practical and Personal Domains in the Curriculum of Teacher Education. *James B. Ingram*

Teaching Method and Academic Success in Sixth Form and University. *J. A. Wankowski*

It may be ordered direct from the Publishers

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada
Agathon Publications Services Inc.
150 Fifth Avenue
New York, N.Y. 10011

May and November
Annual Subscription £2.00 including postage

RESEARCH IN HIGHER EDUCATION
Editor: Charles F. Elton
Department of Higher Education
University of Kentucky

Founded to meet the needs of those concerned with the functioning of two-year and four-year colleges, this journal publishes quantitative studies that contribute to an increased understanding of an institution, or that allow comparison between institutions, that aid faculty and administrators in making informed decisions about current or future operations, and that improve the efficiency or effectiveness of institutions.

Volume 2 (1974), 4 issues, $20.00
Volume 1 (1973), 4 issues, $20.00
(Special rates for personal subscriptions, $12.00 per volume) Additional postage for subscriptions outside U.S., $2.00.

Send orders to: **APS Publications / 150 Fifth Ave.**
New York / New York 10011
U.S.A.

RESEARCH IN EDUCATION

November 1973 *Number 10*

A limited empirical study of the attitudes of serving teachers to educational theory: PHILIP H. TAYLOR

Coalition formation in colleges of education and the introduction of the B.Ed.: J. NIAS

Adjustment in the classroom 1. Teachers' ratings of deviant behaviour: J. A. WILSON

Role orientations and sub-cultures among undergraduate students: LOUIS COHEN and DEREK TOOMEY

The effects of pre-test difficulty on post-test performance following self-instruction: JAMES HARTLEY

University course-changers and potential course-changers: MICHAEL P. HORNSBY-SMITH, CELIA J. NEWBERY and SALLY R. HART

Reviews
by P. E. Vernon, James Drever, Neil Warren, Philip Williams, William Lamont, Louis Cohen, K. Lovell and Ernest Rudd.

Published twice-yearly, May and November

MANCHESTER UNIVERSITY PRESS

NFER
COMPUTER SERVICE

The National Foundation for Educational Research has now installed its own computer—an IBM 1130.

We anticipate a certain amount of spare time in the next year or so, and are prepared to make the computer available to Educational Research Workers at favourable rates.

Card-punching will also be undertaken, at competitive rates.

Details

A comprehensive range of programs is available. These include regression analysis, factor analysis (up to 100 variables), a general tabulation program, χ^2 analysis, and analysis of variance. In addition, there is a wide range of data processing programs for those researchers who need to handle a large amount of data.

Rates

Computer: £20 per hour.

Data Processing: £24 per 1,000 cards. Punched and Verified.

Specimen Costs:
Regression analysis of 8 variables on 40 subjects—£5: 15 tables from 2,600 subjects—£10.

All charges subject to 10% VAT.

Telephone Slough 28161, or write to
A. N. James,
NFER, The Mere, Upton Park, Slough, Bucks.

NEW

Tests of Proficiency in English

- **Assess** 7- to 11-year-olds' attainments in English. Developed by NFER particularly for immigrant children.
- **Test** reading, writing, listening and speaking independently to diagnose the exact area of a child's weakness.
- **Monitor** the child's progress and show when he is ready to join normal junior school classes.

Please send me inspection copies of Tests of Proficiency in English:

Name ....

School ....

Address ....

....

GI
NN

GINN TEST SERVICES, Elsinore House,
Buckingham St., Aylesbury,
Bucks HP20 2NQ

585

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £2·00 post free. Overseas (excluding USA*): £2·30 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: 75p post free. Overseas*: 85p post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Enquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Bucks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the test, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or
2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© *National Foundation for Educational Research in England and Wales, 1974*

Designed by H. A. Senter

Printed by King, Thorne & Stace, Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

14 8/75

In this issue

Moral and religious education
'O' level and CSE examinations
Curriculum and creativity

[illegible] 7/8/75

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd.

EDUCATIONAL RESEARCH

Volume 16 Number 3
June 1974

Contents

Small group teaching methods in higher education *J. P. Powell* 163

Reading standards in Aberdeen, 1962–72 *J. Nisbet, J. Welsh and J. Watt* 172

Reading ages and NFER reading tests *D. Vincent* 176

Elements of elaboration in working class writing *J. Rushton and G. Young* 181

Social class and educational research *T. S. Robertson* 189

Objectives in the humanities *G. Chanan* 198

Establishing children's learning as the criterion for teacher effectiveness *K. B. Start* 206

Conceptions of school among pupils affected by the raising of the school leaving age
A. Smithers, G. Avis and D. Lobley 210

Interpreting research: sampling *B. H. Choppin* 218

The utility of taking notes as an aid to learning *M. J. A. Howe* 222

Book reviews 228

No. [illegible]
[illegible]
[illegible] Edul Psy [illegible]

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 16 Number 3
June 1974

Small Group Teaching Methods in Higher Education

J. P. Powell, University of Papua and New Guinea

Summary
The proceedings of thirteen undergraduate tutorials in a variety of disciplines were tape-recorded and analysed to determine the amount of speech contributed by each member and the cognitive content of the verbal interaction. A small number of groups worked without a tutor. Tutors spoke for rather more than half of the time and when the tutor was absent many students doubled their own contributions and participation was spread more evenly. There was a marked emphasis on providing information in almost all of the meetings but in the leaderless groups this tended to diminish and there was rather more stress on argument. It is suggested that the use of leaderless tutorials be more widely adopted.

Introduction

The application of the techniques and research findings of social psychology to the problems of interpersonal relations and learning in small instructional groups is still very much in its infancy and, although much of the enormous amount of work devoted to the study of group dynamics is of interest to teachers, remarkably little of it has been aimed at increasing our understanding of what goes on in classrooms. A recent textbook (Shaw, 1971) is typical in this regard in that its subject index contains no reference to education, school, college, learning or teaching. Anyone interested in the study of small-group teaching at the university level is therefore thrown largely upon his own resources, although there are some guides to the scanty literature (Powell, 1971) and one excellent discussion of research findings (Abercrombie, 1971).

My own interest in the study of small group teaching methods sprang largely from the seminal work of Mrs Abercrombie (1960) which prompted me to carry out some small-scale experiments with leaderless tutorial groups, the results of which (Powell, 1964) suggested that groups of this type might achieve some educational values which tend to be neglected in conventional tutorials. An early review (Bass, 1950) of the leaderless discussion technique contains no mention of its use for educational purposes and subsequent applications have been overwhelmingly concerned with leadership, personnel selection and psychotherapy.

The study reported here had two main aims: to collect data relating to verbal participation in 'conventional' tutor-led groups and to examine the feasibility of using leaderless groups in university teaching.

Procedure

Many problems confront the investigator of university teaching methods if he wishes to tape-record class proceedings. Both staff and students tend to be wary of having their every word preserved on tape: the wife of a colleague who was in no way involved in the experiment once remarked to me: 'Are you still doing your eaves-

---

Acknowledgements
The study reported here was made possible by a generous grant from the Australian Vice-Chancellors' Committee's Steering Committee for Research and Experiment into Education Matters and I would like to thank the Committee for its support. I am also greatly indebted to all those colleagues and students who participated in the project and to my research assistants, especially Noel Allen and Annette Eassie.

dropping?' Despite careful explanation of the purposes of my work I sometimes failed to persuade students to take part or to dispel the perfectly understandable uneasiness of some tutors. Students who agreed to work without a tutor were always given an assurance that the procedure would be abandoned if they became unhappy with it.

For a number of reasons I was obliged to take whatever groups I could get and this meant that it was impossible to obtain anything which looked remotely like a 'sample', a 'control group' or a 'sound experimental design'. If students and tutor agreed to the tape-recording and they were able to meet at times which did not clash with other recording sessions then they became part of the investigation. Few members of staff were willing to give up their teaching role entirely so the number of totally leaderless groups was always very small and drawn either from my own classes or from Preliminary Year students completing the university's matriculation requirements. Almost all the students were male Papuans and New Guineans. After a number of technical problems had been overcome, recordings of an adequate quality for analytic purposes were obtained from 13 groups. Each group consisted of between seven and eleven students.

Tutors and students were not given any instructions as to how the meetings were to be conducted although it was suggested to the leaderless groups that they could appoint a chairman if they wished. Only the students were present at the leaderless meetings. Most of the leaderless meetings had a definite topic to discuss which was prescribed by the tutor, they were also usually provided with a handout containing some questions or other stimulus to discussion. An interesting question, not taken up in this study, concerns the relationship between the stimulus materials which students are given in tutorials and the amount and quality of the subsequent verbal interaction. As far as I am aware, this important problem has yet to be tackled by research workers since no findings which bear upon it have so far appeared in the literature.

All meetings lasted 45-55 minutes and took place weekly, except for the alternating tutored/leaderless groups which met twice a week, yielding a total of seven to ten recorded sessions with each group. It is not possible to gauge accurately the effects of tape-recording: I feel that most participants become accustomed to it quite quickly but I know from remarks made by students that even at the end of a semester some of them were still inhibited by the procedure. Students were assured that nobody but myself would listen to the leaderless tape-recordings. The only alternative is hidden microphones and that would be ruled out on moral and professional grounds. Records were kept of seating positions for each meeting and at the final gathering for each series all participants (except tutors) were given a sociometric test and a questionnaire. I also discussed the whole project with them at that meeting. This yielded an immense amount of data and so far only a small proportion of this has been analysed in order to establish the amount of verbal contribution from each member and the cognitive characteristics of what was said.

The participation rate for each member was obtained by the use of a technique developed some ten years ago (Powell and Jackson, 1964) and which requires only a typewriter, a metronome and a tape-recorder. Each group member is assigned a code-letter and the appropriate typewriter key is struck at one second intervals according to who is speaking. Blanks are left for periods of silence. In this way the verbal interaction of an entire tutorial can be sequentially displayed on a single sheet of foolscap paper. Each member's participation score can then be calculated as a percentage of the total amount of speech and a mean percentage score can then be derived from these to indicate each member's level of verbal participation over the entire series of meetings.

The content of what was said in tutorials was analysed in terms of a category system (see Table 1) devised to take account of the major forms of cognitive activity which might be thought to possess educational significance at the undergraduate level. There are a large number of category systems currently in use for the analysis of classroom interaction but none of these appears to reflect the pedagogical interests of university teachers. Many of them are concerned with affective responses such as 'shows support' or 'shows hostility' while others concentrate heavily upon the moves made by teachers in initiating interaction. After a great many modifications the system shown in Table 1 was adopted: it is far from satisfactory but most university teachers will, I hope, agree that it identifies cognitive activities which may be seen in all tutorials and which are of major educational significance. The most difficult to define in operational terms, and in my view the most important, is category A, 'arguing'. One of the central skills to be acquired by a student is the ability to develop an argument, to criticize the views advanced by others, to defend his own position, to muster reasons and justifications for his own beliefs, interpretations and claims to knowledge. A number of people worked on the calculation of participation rates but in order to avoid the risk of inter-coder unreliability all the category scores were obtained by one assistant after a number of training sessions with myself in order to clarify the boundaries between the categories. Unfortunately, the poor quality of many of the early recordings limited the number of groups to which the category system could be applied.

Table 1: Category system for cognitive activities in discussion groups

O	Giving an Opinion: Expressing an unsupported belief, value judgment of interpretation. Agreeing or disagreeing. Expressing puzzlement.
I	Giving Information: Providing facts or data. Citing others or books. Recounting personal or group experiences. Reading from text or notes. Summing up.
A	Arguing: Opinion supported by information (O+I=A). Criticizing views of others. Defending own position. Raising objections. Giving reasons or justifications. Developing a point of view.
Q	Asking for Information: Questions eliciting information, clarification, opinions, explanations. (N.B.: Many apparent questions are often objections and should be scored A.)
C	Clarifying: Rephrasing earlier statements. Giving examples. Elucidating quotations from or remarks of others. Explaining. Defining. 'Mulling over' the material.
P	Formulating Problems: Suggesting, proposing, analysing or identifying the problem. Analysis or clarification of problem, suggestions, proposals or questions concerning how to attack the problem. (N.B.: Exclude organizational proposals such as: 'Let's split into three groups to do this.')
G	Group Processes: Proposals for, questions about, discussions of, criticisms of group organization, individual contributions. Expressions of friendliness or hostility towards members. (Include management questions or instructions, e.g. details of assignments, course planning and content. Include laughter and group noise—e.g. all talking or shouting at once.)

Results

It is well known to students, but not to many tutors, that most of the talking in tutorials is done by the tutor. Of the thirteen tutors for whom we have results their mean percentage participation score was 58·0 per cent and the range was from 27·8 per cent to 74·0 per cent. It is clear from this that groups where no tutor is present have available to them, even if they are unable to make good use of it, almost double the amount of time for verbal interaction available to students in tutor-led groups. This can be seen in the data for three of the groups which had alternating tutored and leaderless meetings (see Tables 2, 3, 4). It is worth noting that in the history of science group there were two students who said nothing whatever during any meeting when the tutor was present but they both contributed when he was absent. In most cases the students increased their participation scores quite considerably when the tutor was absent: in some cases they said nearly five times as much, e.g. student number one in both the history and history of science groups. In a few cases students spoke less under the leaderless condition. It would be most interesting to know the explanation for very large increases in contributions from particular students when the tutor was not present but there are too many uncontrolled variables beneath the present data even to speculate about this. One important generalization which is supported by these results is that as tutors tend to talk for about half of any given tutorial period, the time available for student contributions is correspondingly reduced. It would seem to follow that if we want to encourage students to discuss then we should keep out of the room!

In most cases the presence of a tutor tends to increase the total amount of speech in a discussion group, presumably because he is less

Table 2: Participation rates in a history group

Student	*Mean percentage of total speech with tutor present*	*Mean percentage of total speech with tutor absent*
1	5·9	24·8
2	13·6	6·0
3	25·9	20·6
4	6·7	13·6
5	5·2	13·6
6	7·1	12·2
7	7·8	9·2

Tutor's mean percentage rate 27·8%.

Table 3: Participation rates in a history of science group

Student	*Mean percentage of total speech with tutor present*	*Mean percentage of total speech with tutor absent*
1	4·4	24·0
2	0·8	0·5
3	0·0	2·5
4	9·7	18·4
5	6·3	5·1
6	7·2	8·2
7	0·9	5·6
8	6·7	17·4
9	3·0	6·2
10	0·0	7·5
11	5·0	4·6

Tutor's mean percentage rate 56·0%.

Table 4: Participation rates in a political science group

Student	*Mean percentage of total speech with tutor present*	*Mean percentage of total speech with tutor absent*
1	2·5	6·3
2	3·3	5·2
3	7·8	21·2
4	5·3	5·9
5	4·8	0·3
6	3·7	28·8
7	3·2	11·5
8	4·8	5·1
9	1·3	10·8
10	0·2	3·1
11	0·1	1·8

Tutor's mean percentage rate 63·0%.

tolerant of long periods of silence and frequently speaks in lengthy continuous bursts (see Table 5). Although taken as a whole tutorials contain about 25 per cent of silence—one of the leaderless groups spent almost half of their time in silence—it is not common to have periods of silence continuing for longer than thirty seconds (see Table 6). Even half a minute appears to be interminable when one is a participant and few tutors are able to stand the strain without intervening in an attempt to get interaction under way again. This is usually self-defeating as students quickly learn that if they hold out for a while the tutor will soon come to the rescue and keep the monologue going. The tutor is also handicapped by professional norms which pressure him to keep talking since he sees his job as largely that of giving instruction. If the students fail to maintain the discussion then at least he feels that he is earning his money if he keeps 'teaching' them. Some tutors are even told by lecturers to use the time to go over the lecture material again so that it is better understood; in other words, to give another lecture. Most teachers are poor listeners: they conceive their instructional role primarily in terms of telling and explaining and this makes it extremely difficult for them to act as just another group member, although admittedly a well-informed one, with no right to consume most of the discussion time. A consequence of this is that students tend to play the part of passive listeners waiting to be told the answers by an authority figure. One obvious advantage of leaderless groups is that they compel students to adopt other roles and encourage them to emancipate themselves from their customary total dependence upon the teacher. Many teachers do not realize the extent to which students depend upon them. Any one who is sceptical about this should try putting a notice on the tutorial room door announcing that he will be absent because of illness. All of the students will go away: it is inconceivable to them that they could do anything without the leadership of their teacher.

It is worth noting that if the average tutor speaks for half of the available time then assuming that there are ten students in the group and that they each speak for roughly the same amount of time that leaves each student with two minutes in which to make his own contribution to the discussion. Over an academic year of thirty weeks each student thus has a total of only one hour in which to practise his verbal skills and to express and defend his own ideas. It is ironic that those who least need the opportunity to

Table 5: Effect of tutor's presence on total speech

Group	*Tutor/Leaderless*	*Total Speech %*
Py/Hst/A	T	70
"	L	56
Educ/225	T	87
"	L	84
Py/History	T	60
"	L	71
Biology/A	T	60
"	L	63
Politics/C	T	82
"	L	73
Educ/C	L	78
Educ/D	L	78
Educ/D	T	80
Educ/E	T	80
Biology/B	T	69
Biology/C	T	76
Geography/A	T	55
Politics/A	T	73
Educ/B	T	76

Table 6: Periods of silence lasting 30-60 secs. and over 60 secs.

Group	*30 secs.*	*60 secs.*
Py/Hst/L/3	1	0
L/4	3	0
L/5	4	1
L/6	0	0
T/8	4	1
L/9	10	1
T/10	0	0
L/11	4	1
T/12	1	1
Py/History/A/Leaderless		
L/1	2	0
L/2	1	0
L/3	3	0
L/4	4	0
Two Education/T	2	0
One Politics/T	1	0
One Biology/T	1	1
One Biology/T	11	3

Table 7: Desirable features of tutorials as seen by tutors

1. All members participate.
2. Students argue among themselves.
3. All students ask questions.
4. All completed required preparation.
5. All enjoyed the meeting.
6. Discussion keeps close to the point.
7. Students gain some understanding.

practise these skills, the tutors, get ten times as much practice as the average student. One wonders for whose benefit tutorials are arranged!

In order to obtain an indication of staff views on what makes for a good tutorial a sample of thirty staff members at the University of Papua and New Guinea were asked to list features of what they would consider as a good tutorial. Some 41 factors were mentioned and a frequency count produced the seven items listed in Table 7. The most frequently mentioned characteristic was widespread participation in discussion, the second was the development of arguments and the third the asking of questions. It should be noted that respondents were not asked to rank their opinions in order of importance: if they had been so asked then the ordering may have been somewhat different, although I am inclined to doubt this. If a tutor cannot get much response from the students then he is likely to view the whole exercise as a failure. On the other hand if he succeeds in stimulating a vigorous discussion he will count this a success even though the content is perhaps of poor quality.

Let us now turn to the content of the discussions. The category system was devised in order to throw light on the general character of what is said in tutorial discussions and it must be admitted that it is a far from perfect instrument and generates more problems than it resolves. Before we can feel at all confident that there are genuine differences in the cognitive activities of groups when the presence of a tutor is the main experimental variable a great deal more work needs to be done over a wide range of subject-areas. The present results, in view of the small sample (Table 8) and the many important factors which it was impossible to control, must be treated very cautiously and viewed as little more than a collection of hints for future lines of research. The third and final stage envisaged in the present project was the evaluation of the quality of contributions in academic discussion groups: this is perhaps the most important part of the study since university teachers are primarily concerned with the content of what students say rather than the amount of it. This task presents formidable theoretical and practical problems which have yet to be cleared up but work is proceeding on these while at the same time a wider sample of groups is being recorded in order to make it possible for firmer conclusions to be reached regarding the cognitive content of discussions.

Table 8: The cognitive content of discussion

		CATEGORIES						
Group	*T/L*	*I*	*O*	*Q*	*A*	*P*	*C*	*G*
Hst/A	*T*	54·5	7·5	5·7	1·0	1·6	11·6	18·1
,,	L	25·4	15·1	6·8	2·1	0·6	14·5	35·5
,,	Tutor only	63·2	1·5	1·1	0·3	1·5	11·2	21·2
Pol/C	T	37·2	18·0	7·9	7·9	0·6	16·8	11·6
,,	L	26·6	35·0	12·4	6·1	0·5	13·9	5·5
,,	Tutor only	46·2	6·2	6·9	2·6	0·8	19·4	17·9
Py/History	L	60·6	11·0	11·2	7·1	1·0	3·0	6·1
Educ/C	L	27·3	17·3	8·5	23·2	3·9	7·8	12·0
Educ/225	T	28·2	32·8	3·7	16·9	2·7	9·3	6·4
,,	L	23·0	41·3	5·0	17·7	3·9	4·9	4·2

It has already been shown that there is both an increase in and a wider distribution of participation in leaderless groups compared with tutored ones. The results shown in Table 8 suggest several other interesting points which are worth comment. Tutors appear to spend much more of their time in providing information than in encouraging students to advance and develop arguments: this may be partly explained by the fact that in those groups where the leaderless condition alternated with tutor presence the tutor was under considerable pressure to give information since it had been agreed at the beginning that difficulties remaining from each leaderless meeting could be cleared up at the next tutored meeting. One suspects, however, that many teachers see themselves primarily as information-givers; and those who do not conceive themselves in this way are probably forced by the structure of the teaching-learning situation into the role of victuallers of information.

Another feature of the data in Table 8 is the difference between the number of opinions which were offered: it may be that tutors try to control the quantity of opinion by challenging it, asking for supporting evidence, pointing out its irrelevance, and so on, whereas a student response to an opinion expressed by a peer would commonly be the offering of yet another opinion. One would expect a considerable variation in scores in this category according to the type of material being discussed: for example, one would predict a low score in mathematics and a high one in politics.

In view of the importance which university teachers attach to argument in discussion groups it is remarkable how little of what was said fell into this category. Once again, one would expect considerable variation according to subject-matter: I have often heard scientists claim that there is little point in having discussions in undergraduate science teaching because the students do not know enough to be able to contribute anything of value. (This has always impressed me as an extremely revealing and interesting assertion but consideration of it will have to await another time.) In philosophy, however, one would expect almost all of tutorial time to be consumed in argument.

Tutor response

Staff members involved in the experiment were asked to give their views on the advantages and disadvantages of leaderless tutorial groups. The main advantages were seen to be the following: they encourage students to be independent and to rely less on the authority of the teacher; they give students confidence in speaking out and presenting their own ideas; they afford an opportunity for the practice of leadership skills; students are more likely to raise questions which worry *them* rather than questions which the tutor deems to be important; students are less reluctant to reveal their ignorance to their peers; they are able to learn at their own pace; there is an increased pressure on students to undertake the preparatory reading required for a tutorial; most students appear to work harder; leaderless groups add variety to the teaching methods which are currently being used.

A number of disadvantages were also frequently mentioned: an ill-informed student might dominate the group; without guidance the topic might not be explored thoroughly; its use reduces staff-student contact and thus makes it more difficult for staff to get to know students and to identify those areas in which they need most help; some students become anxious if they cannot see the direction the course is supposed to take and if there is no expert present to correct their mistakes; a dynamic leader is necessary for success.

A number of tutors saw considerable value in alternating leaderless groups with tutored ones (as had been done with some of the experimental groups) or in arranging for the tutor to join the group for the final twenty minutes in order to clear up difficulties.

Student response

Many of the above points were also mentioned by students and in addition they listed a number of other considerations. Among the advantages were the following: the atmosphere is more relaxed and students feel able to speak more frankly; there is more time so everyone contributes; it gets students away from spoon-feeding and encourages independence; students learn to help one another; students can learn more from their peers and so gain a deeper understanding; the discussions are more stimulating.

The disadvantages included: we cannot tell whether what we say is correct; much of what is said is irrelevant; one student sometimes dominates the meeting; we make more mistakes than when the tutor is there; it is a waste of time if the preparation has not been done; students sometimes fool around and attendance is poorer.

There are circumstances under which students may well learn more from their peers than from a tutor since the students are all closer to the problem. A tutor often experiences difficulty in grasping just what it is that is puzzling a student since that particular problem ceased to baffle him some years ago. Students also use language which is better understood by their peers. There is also the very important point stressed in the work of Mrs. Abercrombie (1960) concerning the gap between the mental structures (schemata) of the students and those of the tutor; the higher degree of congruance between those of the students could be expected to favour the grasping of explanations and interpretations offered by peers.

As a number of the students pointed out, the atmosphere of leaderless groups is also likely to be more conducive to learning and understanding. All the members are equals in the absence of the dominating and assessive authority figure of the tutor and this cannot but facilitate a readiness to admit ignorance and lack of understanding as well as a preparedness to challenge what is said and a great willingness to seek credentials for claims to knowledge.

Students were also asked whether they wanted more leaderless tutorials in future and if they thought they learnt more in them than in conventional tutorials: 61 per cent said that they did not want any more participation in leaderless groups but 84 per cent felt that they learnt more when the tutor was absent!

Evaluation

It is notoriously difficult to evaluatively compare different teaching methods. Whether or not the students were justified in claiming that they learnt more in leaderless groups would be very hard to determine. There is nothing in the data which really bears upon this point except whatever can be gleaned from the distribution of scores within the categories. Evidence from elsewhere is hard to find. A recent comparison (Clement, 1971) of the retention rates of student-led and tutor-led discussion groups showed that where complex material is to be learnt student-led groups are superior. Clement suggests that this is because of the greater number of student responses and the increased likelihood of individualized encoding (a better fit with schemata in Mrs. Abercrombie's terms) of the material. Another study (Webb and Grib, 1967) demonstrated an improvement in student attitudes and satisfactions for those in the student-led groups without any loss of performance in examinations.

One of the major obstacles to evaluating the effectiveness of small-group teaching methods is that of devising criterion measures which are sensitive to the kinds of student developments which might be expected to result from their use. Examinations are crude but useful measures of the extent to which students have assimilated bodies of complex knowledge but they are of little help in detecting the kinds of learning which we hope discussion groups will foster. This general problem remains unresolved but it is worth pointing out that there is an effective reply to be made to those who refuse to try a new teaching technique until it can be demonstrated that it is superior to methods currently in use, namely, that there is no evidence that the methods currently in use are more effective than any others or, indeed, that they are effective at all!

Conclusions

First, I want to suggest that the major justification for the type of investigation described in this paper, quite apart from any 'results' that may be obtained, may well be that it provides a stimulus to general reflection upon teaching methods and encourages in both staff and students a greater willingness to experiment with new approaches to common problems. The greatest value of experiments with teaching methods may thus lie in their making available a greater range of models from which teachers can choose those which appear most suited to their own subject and personality. My own experience of university teachers in Britain, Canada and Australasia has been that they very rarely discuss difficulties which they encounter in their teaching and if anyone is prepared to collaborate with them in trying out new teaching and learning procedures then this in itself will generate invaluable discussion of the central professional tasks which they face almost every day.

Secondly, I am very sceptical of the value of large-scale studies of university teaching methods which attempt to establish that one method is superior to another. These have produced very meagre results in the past mainly because they have ignored many major variables such as the teaching and learning styles of the participants and, more importantly, because they have been carried out in the absence of adequate criteria for the success of teaching and learning. It is now very clear that we need to pay much more attention to the fine-grain of teaching-learning transactions if we are ever to understand the highly personal and idiosyncratic character of what goes on in classrooms. Such work makes very heavy demands upon the researcher's time and upon his resources of tact, sympathy and perceptiveness but the dividends are likely to be substantial as the work of Parlett and King (1971) has shown.

Finally, I am convinced that we are now justified in encouraging a much more widespread use of the leaderless group technique in higher education, although students do need some initial training in order to be able to exploit its potentialities to the full. This kind of group environment affords students an opportunity to practice a variety of social and intellectual skills which are of the greatest educational significance: the articulation, presentation and defence of arguments; the complex cluster of skills and attitudes which we associate with effective leadership; the willingness to admit frankly to ignorance and faulty understanding; the ability to criticize the views of others without giving offence and to accept criticism with a good grace; engagement in serious discussion with peers without feeling the need to defer to an authority figure; the making of independent judgements which one is able to support with some degree of adequacy and confidence; the possibility of arriving at a deeper understanding of what is involved in working co-operatively and productively with others on challenging intellectual tasks of common interest.

The lecture hall offers almost no opportunities for learning to take place in any of these areas and the teacher-led tutorial, although an excellent technique in many ways, has the disadvantage of imposing a number of major constraints upon students. We now need to be much more confident and daring in encouraging students to take greater responsibility for their own learning.

References

ABERCROMBIE, M. L. J. (1960). *The Anatomy of Judgment*. London: Hutchinson.

ABERCROMBIE, M. L. J. (1971). *Aims and Techniques of Group Teaching* (2nd ed.). London: Society for Research into Higher Education.

BASS, B. M. (1950). 'The leaderless group discussion technique', *Personnel Psychology*, 3, 17–32.

CLEMENT, D. E. (1971). 'Learning and retention in student-led discussion groups', *J. Soc. Psychol.*, 84, 279–81.

PARLETT, M. R. and KING, J. G. (1971). *Concentrated Study: A Pedagogic Innovation Observed*. London: Society for Research into Higher Education.

POWELL, J. P. (1964). 'Tutorials without tutors', *Vestes*, 7, 207–10.

POWELL, J. P. (1971). *Universities and University Education: A Select Bibliography. Volume 2: 1965-70*. Slough: NFER.

POWELL, J. P. and JACKSON, P. (1964). 'A note on a simplified technique for recording group interaction', *Human Relations*, 17, 289–91.

SHAW, M. E. (1971). *Group Dynamics: The Psychology of Small Group Behaviour*. New York: McGraw-Hill.

WEBB, N. J. and GRIB, T. F. (1967). *Teaching Process as a Learning Experience: The Experimental Use of Student-led Discussion Groups*. Final Report, Project No. 5-0923-2-10-1. Washington, D.C.: US Department of Health, Education and Welfare.

Education 3-13

A new journal for Primary Education

In the April issue

Headteachers
Do we need Headteachers?
Aims, organisation and planning: an account of practice in one Middle School
Talking to Headteachers
A School re-organises
Plowden aims: a review

It may be ordered direct from the Publishers

In Britain

The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

Annual subscription £1.00 including postage

In the U.S.A. and Canada

Agathon Publications Services Inc.
150 Fifth Avenue
New York, N.Y. 10011

Reading Standards in Aberdeen 1962-72

John Nisbet, Jennifer Welsh and Joyce Watt,
Department of Education, University of Aberdeen

Summary

The 1962 reading survey in Aberdeen was repeated in 1972, with 99 per cent coverage of year-groups (over 2,500) of 8-year old and 11-year-old pupils. Reading comprehension tests used were NFER Sentence Reading Test AD and Test NS6. The 1972 averages on these tests were within one point of score of the 1962 averages: at age eight, the average was slightly but not significantly higher in 1972; at age 11, it was slightly and significantly lower. The children whose fathers were in semi-skilled or unskilled jobs showed a particularly sharp decline in average performance over the ten-year interval, whereas the standard of performance had improved or been maintained among children from professional homes.

Acknowledgements

The 1972 survey was aided by a grant of £157 from the Scottish Education Department towards the cost of the tests. The Scottish Education Department was not otherwise involved and is not in any way responsible either for the survey or for this report. The 1962 survey was financed by the American Association for Aid to Crippled Children, as part of a larger research study. Dr A. E. G. Pilliner and Miss N. Crawford of the Godfrey Thomson Unit for Academic Assessment in the University of Edinburgh calculated the age allowances and the standardization. Mr G. Horobin and Mr M. Samphier of the Medical Sociology Research Unit in the University of Aberdeen re-analysed all the 1962 data. Mr D. E. Shanks of the Department of Education of the University of Aberdeen gave valuable help in the analysis of data. Students in the Department of Education in the University of Aberdeen assisted in marking and coding. Mrs M. Boyd and Mrs M. Reid of the University Department of Education were responsible foı typing and duplicated material. Mr J. R. Clark, Director of Education, gave permission for the inquiry and supported it with his interest. The highly successful coverage in testing is evidence of the generous assistance and interest of the teachers and headteachers in Aberdeen City, without whose help the survey would not have been possible.

Introduction

The NFER report, *The Trend of Reading Standards* (Start and Wells, 1972), aroused public concern by its finding that the consistent improvement in reading comprehension from 1948 to 1964 had not been maintained in recent years. Evidence from one of the tests used indicated 'a high probability that reading comprehension standards of juniors have declined somewhat since 1964'. Difficulties of organization in the NFER survey, however, had resulted in an incomplete coverage of the national sample. Within the city of Aberdeen, a reading survey in 1962 had covered over 99 per cent of the primary school population; and thus there was an opportunity to conduct a parallel study, within one educational authority, by applying the same tests to complete year-groups of primary school children in 1972. The results proved to be strikingly consistent with those of the NFER study.

Coverage

Pupils in Primary 3 (P3) and Primary 6 (P6) classes in all schools in the city of Aberdeen, including primary departments of direct-grant schools but excluding one special school for handicapped children, were tested in December 1962 and in June 1972. There are no private schools within the city, and the number of children who attend private schools outside the city is probably less than ten in each age-group. On both occasions, the coverage by testing was virtually complete, only the long-term absentees being missed (Table 1).

Tests used

The tests used were: for P3, NFER Reading Test AD (formerly Sentence Reading Test 1); for P6, the restricted NFER National Survey Test NS6. Both tests require the pupil to read and understand a series of sentences: one key word is

Table 1: **Numbers of children tested in 1962 and 1972**

	1962		1972	
	P3	*P6*	*P3*	*P6*
No. on roll	2,866	2,570	2,694	2,735
No. tested	2,865	2,564	2,671	2,716
% coverage	99·9	99·8	99·1	99·3

missing in each, and a choice of words is offered to complete the sense. They are not intelligence tests: if the pupil can read and comprehend the sentence and the alternatives, the choice is obvious. The maximum possible score in Test AD is 35, and in NS6 is 60.

Results

The average scores on the two occasions of testing are given in Table 2. These averages cannot be compared directly, for the first testing was done in the month of December and the second in June. As a result the 1962 groups were two months younger on average than the 1972 groups. The average ages were: P3 1962, 7 years 8·62 months; P3 1972, 7y. 10·96m.; P6 1962, 10y. 8·83m.; P6 1972, 10y. 10·73m. It was necessary therefore to adjust the scores to take account of this difference in ages, and the correction was based on the actual ages of the children, not on the length of time they had been in attendance at school. Possibly this may give the 1972 children a slight advantage. The 1972 test was six months later in the year than the 1962 test, but the average age in 1972 was only about two months higher. The reason is that, when the 1962 children entered school, there were three admission dates in the year (August, January, April) for children who had reached age five by the date of admission; but in 1972 there was only one date (August) covering a younger range, namely those children whose fifth birthdays occurred between the previous March and the following February.

Table 2: **Mean scores of Primary 3 and Primary 6 children in 1962 and 1972 (standard deviation in brackets)**

Year	*P3 average*	*P6 average*
1962	14·42 (sd 7·32)	29·70 (sd 10·38)
1972	15·44 (sd 8·13)	29·47 (sd 11·39)

Two methods of compensating for this age difference were applied, and both gave consistent results. First, an age allowance was calculated: this was the regression coefficient of score on age, which gives an estimate of the increase in average score for each month of age over a twelve-month range. The use of a twelve-month range excluded the risk of distortion from accelerated bright youngsters, or older duller pupils possibly repeating a year; the actual ranges used in this part of the calculation were: P3 1962, 7:3 to 8:5; P3 1972, 7:5 to 8:4; P6 1962, 10:3 to 11:2; P6 1972, 10:5 to 11:4; and in every group the numbers exceeded 2,500.

1. A rough method of correcting for age was to add to the 1962 average an age-allowance equivalent to the difference in ages. For P3, the age-allowance is 0·37 units of raw score per month: the correction figure is 0·86. Thus, for P3, if the 1962 group had been the same age as the 1972 group, their average would have been about 15·28. When this is compared with the 1972 average of 15·44, it is evident that the two averages are very close: the difference is not statistically significant, even on these large numbers. Similar calculations for boys and girls separately gave the same result.

For P6, the age-allowance is 0·23 units of raw score per month: the correction figure is 0·43. Thus, for P6, if the 1962 group had been the same age as the 1972 group, their average would have been 30·13. The difference between this and the 1972 average of 29·47 is statistically significant at the ·05 level. Again, calculations for boys and girls separately gave the same significant difference in favour of the 1962 group.

2. A more precise method is to impose a standardisation by age on the 1962 and 1972 results combined, and then to calculate the scores of the 1962 and 1972 groups separately, using this common standardisation. The scores were standardised to a mean of 100 and standard deviation of 15, and the mean scores were:

P3, 1962, 99·88; 1972, 100·12
P6, 1962, 100·48; 1972, 99·61

The two methods thus give consistent results.

Sex differences

Table 3 sets out the mean scores for boys and

Table 3: Mean scores for boys and girls separately (standard deviation in brackets)

Year	P3 average		P6 average	
	Boys	Girls	Boys	Girls
1962	14·22 (7·70)	14·78 (6·94)	30·48 (10·49)	29·38 (10·22)
1972	14·85 (8·41)	16·07 (7·77)	30·02 (11·59)	28·92 (11·17)

girls separately. At P3, in both 1962 and 1972, the means for girls are significantly higher than for boys; whereas at P6, in both 1962 and 1972, the means for boys are significantly higher.

Social class differences

At the time of testing, under the guidance of their teachers, pupils completed a simple form which included information on father's occupation. This information was available for all except 200 P3 children, and for all except 163 P6 children in 1972. The 1962 data were slightly less complete: information on father's occupation was not available for 304 P3 children and 237 P6 children. As the coverage on this aspect for 1962 was thus 89·4 per cent at P3, and 90·8 per cent at P6, a check was applied to ensure that there had been no selective loss. The mean scores of these reduced groups were calculated. The results (P3, 14·49; P6, 29·86) proved not significantly different from the means for the total group.

The data on father's occupation were coded on the Registrar-General's scale, separating social class III into IIIa (non-manual) and III (manual). The mean scores for each social class separately are set out in Table 4. To allow comparisons between 1962 and 1972, the correction figure for age difference has been added to the 1962 means in the 'adjusted' rows. The familiar social class gradient appears in all the sets of mean scores.

Before a comparison of 1962 and 1972 within social classes could be made, it was necessary to check that the overall social class distribution had not changed markedly.

If a substantial proportion of the non-manual groups had moved out of the city into the rapidly growing suburban housing areas, the comparison of 1962 and 1972 would have been invalid. In fact, this fear proved unfounded: the non-manual proportion in the population had actually increased slightly over the period. For P3, the non-manual/manual percentage changed from 32:68 in 1962 to 33:67 in 1972; for P6, the change was from 33:67 in 1962 to 37:63 in 1972.

The point of special interest is the difference between the adjusted 1962 means and the 1972 means, for the various social class groupings separately. For P3, the non-manual groups (I/II, IIa) have improved their reading performance over the ten-year interval, while the manual groups (III, IV/V) have deteriorated; but none of the differences are significant. For P6, group I/II has again improved, but not significantly; group IIIa show a decline which is significant at the ·05 level; while the drop in average standard for groups III and IV/V is significant at the ·01 level. By far the largest decline is that of group IV/V at P6, a drop of 2.94 points of score, which (on the age-allowance as calculated here) is equivalent to nearly 13 months in reading age.

The size of the gap between the means for social class I/II and IV/V has increased, both at P3 and P6. At P3, the gap has increased from 6.26 in 1962 to 7·41 in 1972 (or, in terms of standard scores, from 0·86 to 0·91). At P6, the increase is much sharper: from 9·62 in 1962 to 13·09 in 1972 (or, in terms of standard scores, from 0·93 to 1·15 sd units).

Table 4: Mean scores by social class (standard deviation and N in brackets)

School class and year	Social class			
	I/II	IIIa	III	IV/V
P3 1962 actual	18·35 (7·0, 505)	16·15 (7·1, 324)	14·16 (7·1, 1,064)	12·09 (6·61, 668)
P3 1962 adjusted	19·21	17·01	15·02	12·95
P3 1972	19·89 (7·3, 533)	17·85 (7·5, 291)	14·90 (7·3, 1,092)	12·48 (7·6, 555)
P6 1962 actual	35·91 (9·4, 471)	33·09 (9·3, 292)	29·16 (9·8, 1,015)	26·29 (9·8, 549)
P6 1962 adjusted	36·34	33·52	29·59	26·72
P6 1972	38·87 (10·0, 618)	31·97 (10·2, 325)	28·01 (10·4, 1,122)	23·78 (10·5, 488)

High and low scorers

Were there more backward readers now than formerly? Or more precocious readers? The larger standard deviations in 1972 demonstrate a wider scatter of scores at both P3 and P6; was this because there were more low scores, or more high scores, or both? The trend of results was the same whichever cut-off mark was used to identify a high or low score. At P3, the proportion of low scorers was approximately the same in 1972 as in 1962 (13 per cent in 1962, 14 per cent in 1972, scoring 5 or less), whereas the proportion of high scorers had increased sharply (8 per cent in 1962, 13 per cent in 1972, scoring above 25). The converse was found for P6. There had been an increase in the proportion of low scorers (8 per cent in 1962, 13 per cent in 1972, scoring below 15), whereas the proportion of high scorers had remained relatively constant (8 per cent in 1962, 9 per cent in 1972, scoring 45 or above).

Conclusions

The NFER findings on recent trends in reading standards are confirmed by this study. While in Aberdeen at the age of eight years, the standard of performance in reading comprehension is relatively unchanged, at age 11 there has been a slight decline in average standard. The difference between the 1962 and 1972 averages, however, is small—only two-thirds of one point of score in a test with 60 items. It would be reasonable to conclude therefore that standards are essentially unchanged over the past ten years.

However, when the results are analysed separately by social class, they cannot be passed over complacently. Whereas among children whose fathers are in professional or managerial jobs, the average standard has improved, or at least been maintained, the picture is different for children whose fathers are in semi-skilled or unskilled jobs. Their average performance, at age 11, is clearly and seriously below the standard of the equivalent social group ten years ago.

The last ten years have seen many changes in primary schools in the city where these surveys were conducted. Especially since the publication in 1965 of the Scottish Education Department memorandum, *Primary Education in Scotland*, there have been more activity methods and group teaching. New schools have been built in new housing areas. Some schools have tried new methods of teaching reading. The examination for selection for secondary education has been discontinued: it was formerly held in November of the Primary 7 year. Whatever the reason, standards have not improved; at the age of 11 the average performance is slightly down; the range of individual differences in reading comprehension is greater than it was formerly, and the 11-year-old children from poorer homes are further behind now than they were ten years ago.

References

Scottish Educational Department (1965). *Primary Education in Scotland.* Edinburgh: HMSO.

Start, K. B. and Wells, B. K. (1972). *The Trend of Reading Standards.* Slough: NFER.

RESEARCH in EDUCATION

May 1974 CONTENTS *Number 11*

Classroom research: a cautionary tale by DAVID HAMILTON, Department of Education, University of Glasgow; and SARA DELAMONT, Bionics Laboratory, University of Edinburgh.

Adjustment in the classroom II: Patterns of adaptation by J. A. WILSON, NICER Research Unit, Queen's University, Belfast.

Teachers' views of the role of Liberal Studies in a Technical College curriculum by MICHAEL ERBEN, Department of Sociology, Garnett College of Education.

Changes in the educational opinions of student teachers associated with college experience and school practice by IVAN HUSSELL, Dunfermline College of Physical Education; and A. SMITHERS, School of Research in Education, University of Bradford.

Are we being fair to political education in the school? by G. MERCER, Department of Sociology, University of Leeds.

Why students wish to reform university government by RICHARD STARTUP, Department of Sociology and Anthropology, University of Swansea.

Book Reviews.

Published twice-yearly, May and November

MANCHESTER UNIVERSITY PRESS

Reading Ages and NFER Reading Tests

Denis Vincent, Guidance and Assessment Service, NFER

Summary

The article outlines some practical and theoretical weaknesses of the concept of 'reading age'. It considers ambiguity in the computational methods for determining the 'average' reading age and the tendency for a reading age to become a fixed property of a pupil rather than an estimate of his standing, relative to time of testing and choice of test. A theoretical standpoint is taken that insufficient is known about the way a reader develops for his attainment to be given a developmental or age-based score. It is further suggested that in any case the relationship between age and reading development is imperfect. This point is developed further when the assumption of a linear pattern in reading development is criticized. It is further suggested that insofar as reading does follow chronological age an age-adjustment in the scale used to express attainment is highly desirable. Practical considerations such as selection of reading material, comparison with other educational tests and the assessment of reading progress, are mentioned. The article concludes with a brief resumé of the problems associated with reading 'quotients' and designating children as 'under-achievers'.

Reading Ages (RAs) and Mental Ages (MAs), together with the accompanying paraphernalia of 'quotients' have had a long history of criticism. Most textbooks on tests and testing (e.g. Vernon, 1960; Cronbach, 1949) criticize and dismiss the use of such age-scales for the expression of attainment scores. More than 40 years ago Thurstone (1926) outlined technically telling arguments for their discontinuation, such as the differences in regression of age on score and score on age and the non-linear relationship between age and test performance. Although such criticisms were not necessarily applied specifically to reading, often they could be legitimately generalized to this field. In spite of this tradition of criticism it is undoubtedly the case that many teachers still think in terms of 'reading ages' (RA) and many tests currently employed do use the RA system.

One of the causes of this persistence of RAs is the inertia associated with the re-usable test constructed in the '30s and '40s such as the Burt, Vernon and Schonell word-recognition tests. These remain in wide use and their age-scale system has become firmly established in teachers' minds as *the* metric of reading attainment.

Average age or average score?

A fundamental problem with the age score is its ambiguity, for there are a number of ways it can be computed. In practice the 'average' raw score for a particular age group is taken as the basis for computation. This average is usually in fact the median—although in many people's minds an average is an arithmetic mean. Both Gulliksen (1950) and Thurstone (1926) have pointed out that it would be possible to find the average age associated with each raw score and use this as an indicator of mental age. The two methods—finding average age for each score group, and finding the average score for each age-group—do not give the same final values, particularly in the case of reading, because age and reading test scores are not perfectly correlated. The basis for calculating reading age is thus somewhat arbitrary. An alternative way of deriving RAs is by calibration with an established criterion for RAs such as the Burt, Vernon or Schonell word-recognition tests. This however would seem to be a retrograde step when the dated nature of the original tests is considered.

There are better scales now available, although it would be naive to expect the use of RAs to cease rapidly in the future: at the NFER's Guidance and Assessment Service in Slough we are often asked for advice on the conversion of standardized test scores into RAs in preference to the system of standardized scores we generally employ. In addition, the observations made by teachers attending our courses on objective testing indicate that RAs—and the more primitive tests with which they are sometimes associated—enjoy strong loyalty amongst teachers.

In spite of the slow rate of change in practice there is a case for stating at greater length than textbooks allow the objections to RAs. If nothing else, it may do something to reassure those who are at present uneasy about the systems such as standardized scores which more recent tests tend to employ.

Reading age as a label

Many of the general objections surround the misleading implications of RAs. To say a child 'has' a particular RA implies some kind of absolute judgement about his development as a reader, and there is a genuine danger of misleading the uninformed when we describe reading attainment in terms of a child 'having' a particular 'reading age'. Something inert or fixed is implied by this in a way which 'standardized score' and 'percentile' avoid. In effect RAs have a tendency to stick to or label the reader—often in a perjorative way.

It is often not understood that a 'reading age' is only an equivalent value given to a raw score and is at best only an estimate of attainment based on performance on a particular test at a particular time. Goodacre (1972) has stressed for a lay readership that a child's attainment is relative and linked closely with the test used. There is a similar need however amongst both teachers and parents to realize that scores can only give estimates of attainment and that these estimates are prone to error from a number of sources. In the better-constructed tests this is expressed statistically as the 'standard error' of the test. It has, however, not been the practice to state standard errors in age-scale reading tests, and in some cases at least it is not clear how this could be done with convenience.

Reading development and chronological age

The distinction between qualitative and quantitative testing also needs to be considered. Although some quantitative value—such as number of words read—may appear to typify an age group, we cannot assume that there is any degree of uniformity within that group in the quality and development of their reading skills; certainly two backward readers of the same RA may have very different reading problems and behave very differently *as readers*.

It would therefore seem fair to reserve the comparison with age for aspects of development which are qualitative and genuinely linked to age in a uniform way. Some aspects of cognitive development, particularly those concerned with conservation, may qualify for description in these terms. Reading—at least in its quantitative aspects—is only associated with age fortuitously. Reading development lacks its Piaget, and until it can be described in the terms in which cognitive-developmentalists describe mental growth it will not have a true age-scale.

This distinction can be further illustrated when we consider the difference between diagnostic and attainment tests. Attainment is the outcome of development and learning, but this development takes place in the psychological equipment and processes used for reading. To observe this side of the reader's functioning requires a more elaborate and qualitative instrument than the usual attainment test. At the least it will require an analysis of errors while at the more elaborate stages it will examine the various motor, perceptual, linguistic and symbolic processes which may be associated with malfunctioning or under-development. Here genuinely age-linked development may be of concern, such as acquisition of left-right differentiation, but the process cannot necessarily be described quantitatively. For this reason we find subtests in Daniels and Diack's *Standard Reading Tests* (1958) which do not possess norms or a means of providing raw scores which in turn have additive age-value in years and months.

Many of the problems that weak readers display are in fact only tenuously related to age, just as the acquisition of facilitators of reading performance is not really related to age. Reading is essentially a *learned* behaviour and although some aspects of the process, such as those associated with 'reading readiness' may have close associations with development and age, one might reasonably expect other temporal factors, such as length of schooling or number of hours spent in reading instruction to be more closely associated.

The reading ages of backward readers

The real victims of the RA system are of course the poor readers, and use of RAs does little to illuminate their problems. A reader may have developed in all the areas necessary for successful reading except one or two which are by themselves, or by interaction with social and personality factors, impairing his performance. In such cases nothing is to be gained by speaking of a 'low RA' which can misleadingly imply some kind of 'immaturity' or 'underdevelopment' of a vague and global nature whereas a thorough investigation will perhaps reveal some particular difficulty. Even worse, the 'low RA' may itself be interpreted as a causal factor in the reader's poor performance.

The situation is worsened when the weak reader is compared with other children. It cannot be assumed on the strength of an attainment test that the quality and development of his reading is like that of younger children for whom his raw score is the norm. There is also the possibility that two readers of similar chronological age (CA) and RA may achieve their similar attainment scores while displaying rather different patterns of development and mastery in the psychological sub-skills of reading. Yet the implication that their development is qualitatively similar is inescapable if they are to be discussed in terms of RAs. Where reading has 'aged' so differently in two people it seems to be perverse to use a single identical value to describe both pupils.

Reading age as a metaphor

The essence of the difficulty is that educational attainment can only be given an 'age' value in metaphorical terms; reading attainment may be empirically associated with age, but the two ideas are logically different. The role of the metaphor is to indicate an identity which can only exist in the poetic imagination. When its use is translated to educational measurement the effect, as we have seen, is to make misleading implications. A thoughtful and aware user of RAs would probably be alert to those limitations, just as he might be aware of the limitations of tests in general; the fact remains though that most other methods of expressing attainment do not have these pitfalls. In any case not all those involved with the child as a reader, whether they are parents or teachers, will appreciate the way the notion of reading age can mislead. In addition there are further assumptions on which RA scales are based which must be discussed. One of the least tenable of these assumptions is that reading progress can be plotted in terms of regular linear growth. Richardson (1941) points out that for a MA unit to be a real unit of measurement the function it represents would have to progress regularly with age. Thus one year's mental growth would be equal to another's. Most of the reading tests giving RAs were constructed on this principle: an equal number of items is selected for each year and each individual item is worth the same amount in terms of months or of reading age. It has been pointed out by Vernon (1960) and Daniels and Diack (1958) that rate of development in reading is probably irregular and characterized by sharp rises and plateaux rather than a smooth line or curve. This irregular progress is found in even the more easily quantified aspects of reading such as word-recognition which is characterized by changes in the percentage increase from age to age (Reed, 1953).

It is true that during the junior years the rapid development of reading attainment is such that it appears for a time to develop with age but at the secondary level there is a marked decline in the increase. A ceiling is reached so that the attainment of fifteen-year-olds and eighteen-year-olds may show little difference when measured on conventional attainment tests. The analogy between reading and age is thus a limited one and depends upon the age at which the initial teaching of reading commences. It also follows that where any changes are made in the age at which reading instruction begins or in the rate at which reading is allowed to develop there will be a substantial discrepancy between CA and RA so that existing norms will no longer apply. The use of the age-unit however implies that somehow these ages *ought* to coincide.

Need for age adjustment

A further problem with age-scores is that for comparisons they place all children on a single line of development. However, in view of the irregularity in age units and percentage improvement between age-groups, comparison between children of different ages is less meaningful than that which can be made within a particular age-group. A child's attainment can be most fairly evaluated only in relation to those who are in other important respects his equals. Chronological age is one of the most important of these respects and is therefore used as the basis for standardizing scores. By contrast an RA score takes no account of CA so that the youngest child in a class whose performance is average for his age may be compared unfavourably with older children whose performance is also average for their age.

Reading ages from standardized scores

It can be argued that when reading material is classified in terms of suitability for particular age groups or indeed reading age groups it is genuinely necessary to know children's RAs. The necessity for a tested RA would however depend upon how precisely the age-level of the book had been calculated. Different tests in any case may give different RAs and the effectiveness of the matching of books and readers probably depends upon the astuteness of the teacher's judgement as much as the reader's test score. Under these imprecise conditions a standardized score would seem to be just as effective a basis for matching readers with books. Where a child's score falls close to the mean value there would seem to be every grounds for taking his CA as an estimate of his 'reading age'. In cases where standardized scores are further from the mean it would suffice to find the chronological age group for which the raw scores are close to the mean value in standardized scores. In practice this involves reading the raw-score conversion table 'backwards': the tables are entered at the row containing the raw score and then read across until the mean standardized score value—usually 100—is found. The age at the head of the column is then taken as the estimated 'age-equivalent'.

Selection of reading material

This procedure is sometimes recommended by the Guidance and Assessment Service to users of reading tests who insist on having some kind of age-equivalent value given to a child's performance, although often a better alternative answer to the enquirer's problem can be found. It is also strongly advised that any age-equivalents found are not recorded or used for official purposes and pointed out to users that a very high or low score may require considerable extrapolation beyond the published tables before a theoretical age-equivalent can be found.

It is probably the case that far more effective means of book-selection are available, although most require more work than simply using an RA. The most promising of these is the use of readability formulae. The New Zealand Council for Educational Research (Elley, 1969) has devised a simple method of classifying reading material into ten levels based on the frequency value of the nouns. A reading test has been constructed which uses this system of levels, so that a reader can be assessed in terms of the level of readability at which he is capable of reading. Such a procedure is very much more scientific than that of 'reading age-suitability'; it ties the reader's capacity to a clearly understood body of content which relates to the nature of the reading material in a direct and calculable way.

This 'criterion-referenced' or 'content-referenced' approach to testing has the advantage of focusing on what the reader can or cannot do, regardless of age, and regardless of what any other reader can do. There are relatively few such tests available, although some of Daniels and Diack's *Standard Reading Tests* employ this principle and tests which have recently been developed at the NFER for assessing literacy and oracy in immigrant children use such a principle. In the latter case performance is described in terms of three levels of mastery, and each level is interpreted in terms of its significance for teaching the child concerned.

Comparability and assessment of progress

Attainment tests are a useful means of comparing performance in different subjects and monitoring progress within a single subject. However RAs do not lend themselves to such purposes as well as standardized scores. Reading is the only area of the curriculum in which the age concept is still applied, so that direct comparisons with Mathematics and English test scores are not possible using an RA scale. A common scale for all such tests is clearly desirable and this would require the expression of reading attainment to conform to the majority practice.

Measurement of progress is in any case a more complicated matter than many teachers assume—educational researchers employ relatively elaborate statistical techniques to determine whether progress or improvement has genuinely occurred—and the use of RAs makes it more so. Tests employing RAs rarely state the amount of statistical error to be attached to test scores so that, within certain limits of confidence, differences between scores could be interpreted as significant.

In addition, the imperfect relation between reading progress and age means that the amount of progress a pupil makes will have different value at different ages. A two-year backward pupil who increases his RA by two years at the age of nine is making more ground, in terms of the proportion of children excelled, than a pupil who does the same at 13 years. This is because a score in terms of years and months takes no account of the distribution of pupils about this point, although it is fairly well-established that the range of scores below the average point becomes increasingly extended with age.

'Quotients' and reading potential

Age-based scores have been widely used to form quotients in the past, and reading quotients (RQs) are still sometimes employed. Strictly speaking an RQ is the result of RA/CA x 100, although standardized scores are sometimes called 'quotients'. In the former case the RQ suffers from the fact that the less than average reader has an increasingly difficult struggle to make his RA equal to his CA with age, so that an RQ of 90 at nine years indicates a lower relative standing than such an RQ at 13 years. It also follows that a pupil may display a decline in his RQs over the years while remaining a better reader than the same proportion of children throughout this period.

A further development of the quotient system was based upon the relationship between attainment and some test score taken to show 'potential', customarily an intelligence test. Here quotients were formed using such formulae as RA/MA x 100 or even RQ/IQ x 100. These 'attainment quotients' have been the subject of controversy in the past (Pidgeon and Yates, 1956; Burt, 1957; Crane, 1959) and it does not appear that the debate has been resolved in a way teachers would find satisfactory. In general a distinction is still made between 'underachievers' and those whose attainment appears to match their measured potential, but there is no reason to suppose that quotients are the only basis on which this distinction can be made. The use of profiles and confidence limits based on standard errors would in fact accord more with text-book practice. This would of course involve the use of some scale other than an age-based one.

The problem of whether there is such a thing as measurable 'reading potential' lies outside the scope of the present article. It is, however, unlikely that continued use of RAs will add anything useful to our practical or experimental knowledge.

Conclusion

The traditional case for retaining reading age units is that they are easily understood and recognised by teachers and parents. One suspects however that they are just as easily *mis*understood. The 'age' metaphor is one that is certainly embedded in common parlance but in view of the many superior methods of expressing attainment now available and the fact that reading is the only area which clings to the age concept there is a strong case for abandoning it in favour of the kinds of scales employed by modern mental and scholastic tests.

References

BURT, C. (1957). 'The relationship between ability and attainment', Bulletin of the National Foundation for Educational Research in England and Wales, No. 9.

CRANE, A. R. (1959). 'An historical and critical account of the accomplishment quotient idea', *Brit. J. Educ. Psychol.*, **29**, 252–9.

CRONBACH, L. J. (1949). *Essentials of Psychological Testing*. Harper & Row: New York.

DANIELS, J. C. and DIACK, H. (1958). *The Standard Reading Tests*. Chatto and Windus: London.

ELLEY, W. B. (1969). 'The assessment of readability by noun frequency counts', *Reading Research Quarterly*, **4**, 411–27.

GOODACRE, E. (1972). 'Reading tests: what they are and why they matter', *Where*, April, 114–7.

GULLIKSEN, H. (1950). *Theory of Mental Tests*. Wiley: New York.

PIDGEON, D. A. and YATES, A. (1956). 'The use of tests in the classroom. The relationship between ability and attainment—an examination of current theory'. The Bulletin of the National Foundation for Educational Research in England and Wales, N.8.

REED, G. F. (1953). 'A reading test for Hull', *Studies in Education*, May, 46–56.

RICHARDSON, M. W. (1941). 'The logic of age scales', *Educational and Psychological Measurement*, **1**, 25–34.

THURSTONE, E. L. (1926). 'The mental age concept', *Psychol. Rev.*, **33**, 268–78.

VERNON, P. E. (1960). *Intelligence and Attainment Tests*. University of London Press: London.

Elements of Elaboration in Working Class Writing

James Rushton, University of Manchester School of Education
and George Young, Bolton College of Education (Technical)

Summary

In an attempt to define the concepts 'elaborate' and 'restricted code' by which the language of middle-class and working-class people has been said to be distinguishable, the authors conducted an experiment in which a small group of public school sixth-formers and a small group of young car-workers of similar age and poor educational record wrote essays on a variety of themes. These were closely analyzed for characteristics of verbal usage. It was found that when essays on technical topics were considered alone, the general linguistic disadvantage of the working-class groups was wiped out. This suggests that areas of working-class language may, after all, be 'elaborate' and complex rather than 'restricted'.

Introduction

The study of the relationship between language and patterns of social stratification has been influenced by a theoretical model first formulated by Bernstein and then later elaborated by him in a succession of papers (1958, 1959, 1960, 1961, 1962a, 1962b, 1971). The basis of the model is the assumption that the widely documented linguistic differences between the middle and working classes result, 'from entirely different modes of speech which are dominant and typical of these strata' (1961).

The dominant mode of speech of the middle class is said to be the 'elaborated' code, a code of abstract and explicit meaning in which the connections of logical thought may be clearly articulated, a code which allows the significance of things to be removed from their particular context and seen in a universal perspective a code which gives the individual an awareness of self, a sense of identity in isolation from that of his group. In contrast, the dominant mode of speech of the working class is said to be the restricted code, a concrete and descriptive code whose meanings are condensed and implicit, a code in which the awareness of self and the refinement of rational thought are sacrificed to the needs of metaphorical power and of the solidarity of the social group.

The implications for education of the hypothesized existence of two distinctive class codes lie in the fact that the elaborated code is seen not only as the dominant mode of speech of the middle class but also as the medium of social control and intellectual exchange in the schools which can alone give access to the orders of meaning of higher learning. Differential access to the elaborated code thus becomes the key variable in the analysis of the under-achievement of working-class children.

In the most general terms, this is Bernstein's model as it was originally formulated. As such, it has not only been given wide currency in educational debate but has also given an impetus to empirical research in this area. The roots of what might be called the generative power of the model lie in the assumption, which Bernstein makes, that the psychological orientations of the two codes are expressed in their different communicative strategies, what he calls their different 'verbal planning' processes, which in turn can be objectified in the *raw frequencies* of their structure. To researchers this raised the hope that the difficult analysis of the many process variables involved in the study of the determinants of educability could be circumvented by the investigation of the more tangible and more easily quantifiable variables of language structure. A great deal of space in Bernstein's early theoretical papers (1958, 1959) was devoted, therefore, to the specification of those linguistic items whose frequency of occurrence in either code could be taken to signify different orientations towards the organization of meaning.

The line of theoretical debate and empirical research which Bernstein initiated has proved vulnerable to attack from a number of directions. Some of the criticisms have been limited and highly specific, concerning, for example, the methods used in the collection, treatment and presentation of data (Coulthard, 1969). Others have been more general and perhaps more provocative, for they have pointed to a lack of clarity in the formulation of the model and consequent weaknesses in the design of the research which has followed it. Of these general criticisms two became motivating factors behind the development of the present research. They are:

(i) that, in studies of social class differences in language usage there has been too great a preoccupation with speech as the mode of communication (Robinson, 1965a);

(ii) that there has been a failure also to allow for the contaminating influence of contextual or situational factors upon all levels of linguistic performance (Robinson, 1965a; Lawton, 1968; Coulthard, 1969; Bernstein, 1971).

References to written language are very rare in Bernstein's work. The model of the two codes thus provides no theoretical framework for the analysis of written data. As evidence of the confusion which has resulted two of the researchers who have ventured into the study of written language have adopted theoretical positions which appear to contradict each other. Lawton (1968) used Vygotsky's (1962) remarks about the autonomous nature of written language to suggest that the act of writing might require a form of verbal planning closer to that of the elaborated code. His prediction that, as a result, social class differences in writing would be greater than those found in speech was confirmed. Robinson (1965a), on the other hand, has argued that as speaking is the skill which requires the more rapidly produced sequence of co-ordinated activity it may be that working-class subjects will find it easier to produce elaborated language in their writing. Certainly, the limited evidence which Robinson offers us suggests that the theory of the two codes may find less strong and consistent support in written data.

'It is certainly the case', Bernstein (1971) has said, of his own work, 'that up to 1968 there are few explicit mentions of the constraints of specific contexts upon the patterning of grammatical and lexical choices'. It is all the more interesting, therefore, that a number of scattered researches have shown (Lawton, 1965; Robinson, 1965a; Henderson, 1970) that, where aspects of the context in which language is produced, particularly the subject for discussion or the eliciting question in speech and the essay topic in writing, are altered systematically, stylistic variation will occur in the language of middle-class subjects and in that of working-class subjects also, although generally to a lesser degree. The findings of these researches are strong enough to throw some suspicion upon the conclusions drawn by those researchers who failed to control for the hidden variables of contextual constraint. It could, after all, be argued that the observed depression in working-class language scores are a consequence not of any failure of competence but of an inability to draw upon that competence in certain situations.

The findings also present Bernstein's theory with a major problem of incorporation. Robinson and Lawton use the term 'code-switching' to describe the variation in the language scores of social class groups across different contexts but their application of this term to the language of working-class subjects is suspect, since in terms of the model as it has come to be generally accepted, such subjects are prevented by definition from 'switching' codes. Bernstein, it appears, has chosen to maintain the integrity of his conception of two class-related codes with distinctive verbal planning processes by making a distinction between the codes and their linguistic variants. It is necessary, he says, 'to make explicit the distinction between the patterning of speech evoked by specific social contexts (speech variants) and the concept of code as a regulative principle controlling speech realizations in diverse social contexts' (1971). The distinction brings with it a limitation of research objectives for, as their underlying regulative principles are not directly accessible to observation, it is the contextual variants of the codes which become the objects of analysis. To identify the codes, say Bernstein, one now needs to, 'obtain samples of speech in critical contexts and examine their syntactic and lexical properties in the light of each context according to the explicitness and specificity of the meaning. On this basis, one could only infer code after examination of speech obtained from selected contexts' (1971).

Bernstein's concession to the influence of context upon language represented the point of departure for the present research. The aim of our investigation was to see whether the previous findings of consistent social class differences in language useage held good when writing and not speech was the mode of communication and a more rigorous control was exercised over the influence of essay topic upon language.

The weight of earlier research led us to expect a massive overall inferiority in the language of our working-class subjects in the use of complex or 'elaborated' structures when compared with that of middle-class subjects of the same age. We did not, however, expect this inferiority to be characteristic of all areas of their language. We deliberately chose to include among the selection of titles which we gave to our subjects certain technical topics with which, we felt, the working-class subjects would be more familiar. On these topics we expected the latter to be able to 'lift' their language performance to an equal level of complexity in certain limited areas of structure. Thus in formulating our objectives we found that we could accept analytically the concept of 'elaboration' in language structure while leaving open the question of whether or not there are such things as distinctive class codes.

Design and procedure

(a) *The samples*

The sample of working-class subjects consisted of 25 adolescents working as operatives in a car factory and attending a college of further education on a day-release basis as part of a general course of industrial training. The middle-class subjects were 22 pupils in the science courses of the sixth form of a major public school. The latter group, however, were very unlikely to be representative of the middle class as a whole. All of the public-school group except one were intending to go to university at the end of their course. All had fathers whose occupations fell into the higher categories of the Registrar General's Index of Occupational Classification (1966). The working-class subjects were without exception the sons of manual workers. All had attended secondary modern schools. Where streaming had been in operation, as it was reported to have been in eighteen out of the twenty-five cases, all except three had finished school in the lower streams. Nor did any of the working-class subjects possess a formal examination qualification of any description. They had all left school at the statutory leaving age and the course they were attending was only one of minimal industrial training. The mean age of this group was 17·1 years; of the middle-class group 17·5. By choosing the groups from these two extremes of social background we hoped to obtain a greater contrast than had been shown in previous work.

The two groups were tested for ability on the AH4, a test divided into two parts, the first testing verbal ability with a numerical component, the second spatial ability. The mean scores for the two social class groups were as in Table 1:

Table 1: Ability of the two class groups

	Working-class	*Middle-class*
AH4, Part 1 (Verbal)	25·6	51·6
AH4, Part 2 (Spatial)	35·9	54·0

Two points about the pattern of these scores require emphasis. Using Raven's Progressive Matrices Test, 1938, and the Mill Hill Vocabulary Scale form 1 Senior, Bernstein (1960) found that among his working-class sample the language scores obtained on the latter were depressed in relation to the higher ranges of the scores on the Matrices test. Although the dichotomy of abilities presented by the separate parts of the test used here is not strictly comparable with that produced by the use of two separate tests, the pattern of scoring we found showed a remarkably similar pattern of relative depression of verbal ability. What is more, the pattern of depression existed at a much lower level of general ability. In Bernstein's study the gap between the social class groups closed considerably on the non-verbal test. In our study the massive superiority of the public-school group was only moderately reduced on the non-verbal section. The large difference between groups on this second section places the marked inferiority of our working-class subjects in verbal skills within a total picture of inferiority in ability as a whole and reinforces the point made earlier in describing their educational backgrounds that with them we were dealing with a less selected and less able working-class group than those of many earlier researches.

(b) *The essay titles*

The subjects were at first given nine essay titles on which to write. The essays were done within the school and college timetables and handed in over a period of some months. The time limit allowed for the writing of each essay was very broad and no situational controls were exercised other than those which usually prevailed in the classrooms concerned when writing was in progress. The essay titles were categorized according to the style of writing and treatment of subject matter which they were intended to evoke. Three categories were produced by the use of this criterion. They were labelled as:

(i) *Imaginative descriptive*

'The City at Daybreak';
'What the world might at first look like to a man who had been blind but had just been given his sight';
'A walk across a disused aerodrome on a hot summer's afternoon'.

(ii) *Opinionative discursive*

'The pleasure of making things';
'Speed: the fascination of danger';
'The world in fifty years' time'.

(iii) *Technical explanatory*

'A workshop lathe: what it looks like and how it works';
'Describe the appearance and function of any piece of machinery, showing how the way it is constructed enables it to perform its function';
'Describe how the way a bicycle is constructed enables it to perform its function'.

Only the last of these titles produced such a small response that it had to be removed from the analysis.

(c) *The operational definition of the hypotheses*

The grammar used in the analysis was the 'Scale and Category' or 'systemic' model developed by Halliday (1961), Halliday, McIntosh and Strevens (1964). Research of this kind requires the selection of certain items for frequency analysis. The model of grammar used provided us with a comprehensive description of English structure which allowed us to pursue our investigation within a systematic conceptual framework but it could not of itself generate the criteria for the selection of items for analysis. To understand the criteria which we used we need to return for a moment to consider the general nature of our predictions.

We predicted, first of all, that we would find a general superiority of the middle-class group on the items selected when the mean scores on all the essays taken together were compared but that this superiority would not be apparent on the technical titles taken separately since the need to express technical concepts would compel the working-class subjects to use certain 'elaborated' items which they might not have used elsewhere. We required, therefore, linguistic items whose frequency of occurrence had been established not only as a distinguishing characteristic of one or other of the two codes or as a discriminating variable in studies of social class differences in language usage but also as an indicator of a 'technical' style of writing. A number of items especially those concerned with the use of the noun and the verb, were central to our thesis. These items and the reasons for their choice are described below.

Bernstein pointed to their limited and rigid use of adjectives as a principal reason for the failure of restricted code users to make elaborated qualifications in their handling of concepts. The adjective, as is commonly known, is a class of word (e.g. *big*, *powerful*, *Czechoslovakian*) whose function is usually to modify the noun in the structure of the nominal group (e.g. *the big man*, *the powerful antidote*). Not all nominal groups are susceptible to modification and consequently open to contrastive analysis (*the big he*, *the powerful it*, as nominal groups would not usually be considered grammatical). For this reason we confined our research to the more modifiable groups, those with nouns or ordinals at their heads (e.g. *the strong man*, *the big two*) and excluded from our analysis certain very common modifiers, the deictics (words like *the*, *no*, *some*) whose numbers threatened to inflate our totals with modification of a simple nature. As indices of special complexity we made a count of those nominal groups which contained both a modifier and a qualifier (e.g. *the huge Colonel of the Guards*, *the costly medicine in the window*) and of those which contained two or more modifiers (e.g. *the big Czechoslovakian Colonel*, *the strong but costly medicine*). The second of these two indices was all the more useful to us because Quirk (1962) had remarked upon a tendency to use heavy nominal modification of this kind in technical writing, a characteristic also noted by

Leech (1966) in his study of advertising copy. From the point of view of both of our hypotheses, therefore, there seemed good reason to choose this area of structure for analysis.

Another suggestion made by Bernstein was that the crudeness of the structure of the verb in the restricted code was a reason for the difficulty found by those limited to the code in the description of processes. Bernstein (1962b) and Lawton (1963, 1964) followed this suggestion up with a count of an item which they called 'complexity of verbal stem'. As a definition of structural complexity this measure has recently been criticised (Coulthard, 1969). We decided, therefore, to use other, less ambitious, indices. One of these we found in the use of the passive voice (constructions such as *The medicine was bought in the shop*, *The Colonel was killed at Austerlitz*) which both Bernstein and Lawton had found to discriminate between social classes and whose absence or low frequency in a text the latter had taken to be 'symptomatic of a limited control over language use' (1968). Since Svartvik (1955) had also found scientific and technical texts to contain a high proportion of passives the item could be taken as a critical variable for both hypotheses.

The two hypotheses can now be given a more detailed operational formulation. It was predicted that:

(a) an overall comparison of the written language structure of the two social class groups would show the middle-class group to use more nominal groups with modifiers, with modifiers and qualifiers, and, where modifiers occurred in the writing, more modification of a complex kind, and a greater number of passives;

but that:

(b) where the use of extensive modification and the passive was concerned, the superiority of the middle-class groups would not be apparent on the technical titles when these were analysed separately.

(d) *Method of scoring and treatment of results*

The occurrences of all items in the nominal group were treated as a percentage of the total number of occurrences of nominal groups for each individual. In the same way, the occurrences of the passive were treated as a percentage of the total number of occurrences of all verbal groups. For each individual percentages were obtained for each item on each of the stylistic categories. The three percentage scores thus obtained were then summed and averaged to obtain the overall essay score for hypothesis (a). For both hypotheses it was considered necessary to partial out the possible influence of verbal ability scores upon language performance by the use of analysis of covariance.

Results

For the prediction of the movement of an individual item to be fully confirmed it was necessary not only for the movement to be in the predicted direction and part of a general pattern of scoring but also for it to be apparent to a degree which would preclude chance expectation. By this criterion hypothesis (a) can be seen to receive moderately strong support from the data in Tables 1 and 2. All item scores move with the hypothesis and two of these to a highly significant degree. The general picture, here as elsewhere, is the already widely documented one of social class differences in language usage, with a pronounced middle-class superiority on certain critical items.

The picture of internal variations in structure presented in Table 2 provides hypothesis (b) with

Table 1: The variation in overall mean percentage scores as a function of parental socio-economic status after covariance adjustments for uncontrolled verbal intelligence scores (AH4, Part 1)

Linguistic Item	*Middle Class*	*Working Class*	*df*	*F Ratio*	*p Value*
Nominal groups with one modifier	19·1	17·6	1·44	2·91	N.S
Nominal groups with a modifier and a qualifier	7·6	4·8	1·44	18·50	0·001
Nominal groups with complex modification	3·2	2·6	1·44	2·02	N.S
Passives	18·3	10·4	1·44	6·70	0·01

Table 2: The variation in mean percentage scores of the two social class groups on those linguistic items selected for analysis on the technical titles as a function of a parental and socio-economic status

	Middle Class			*Working Class*					
Linguistic Item	Imaginative Descriptive	Technical Explanatory	Opinionative Discursive	Imaginative Descriptive	Technical Explanatory	Opinionative Discursive	*df*	*F Ratio*	*p Value*
Nominal groups with complex modification	3·8	3·1	2·8	2·5	4·0	1·3	1·44	0·48	N.S
Passives	9·2	33·8	12·0	5·3	21·1	4·7	1·44	24·81	0·001

rather more uneven support, with one item (Passives) moving dramatically against the hypothesis and the other moving equally dramatically in its favour. Even so, the fact that the working-class group were able, not only to equal the percentage of nominal groups with complex modification produced by the middle-class group on the technical titles, but also to achieve a numerical superiority sufficient in itself to wipe out the social class differential on this item predicted on the overall comparison, must lead us to place the greater emphasis on the movement of this item. The second hypothesis, we feel, is given tentative, if somewhat scattered, support.

Discussion

At the outset the purpose of this research was seen as an attempt to test Bernstein's hypothesis of two separable and structurally defined codes on its strongest ground, in the form of a blanket comparison of the written language performance of two highly contrasted social class groups, and on what appeared at the time to be its weakest, in the form of a study of some of the internal variations in language scores which are left unexplored in a blanket comparison.

The items which we extracted for analysis and whose movements were reported in this paper were clearly few in number. Equally clearly the impressive movement of one of them (nominal groups with complex modification) in support of our hypotheses does not make a whole case. Yet, the finding that, for example, a sample of working-class adolescents, without formal educational qualifications and with an almost uniformly depressed verbal intelligence, score, were able, as a group, to produce in their technical writing, a higher frequency of complex nominal modification than were a sample drawn from the science sixth of a major public school, is, in the perspective of past research in this area, an extraordinary one and one which raises a number of questions about the theories which provoked our investigations. There are grounds, we feel, for further exploration.

Labov (1969) has suggested that the most useful service which linguists can perform at the present moment is to clear away the illusion of the 'verbal deprivation' of classes and sub-cultures. Now it is not our contention that the ascription to the language of our working-class sample of the phrase 'less elaborated' is undeserved. The confirmation of our first hypothesis shows this, conclusively, not to be the case. What we do suggest is that it is premature as yet to talk of the 'restricted' or 'impoverished' nature of working-class language without further qualification. Once we accept that the language produced by the middle-class group was the more 'elaborated', and our hypotheses went no further than this, then we must also admit that, in achieving higher frequencies on a particularly critical item, the working-class group were also enjoying a measure of access to elements of elaboration. Such an admission would seem to do some damage to hopes either of establishing a structural definition of the elaborated and restricted codes or of finding a social class location for them. It also makes Bernstein's (1971) concession to the influence of context upon linguistic behaviour seem a none too premature one. Choice of topic or essay title does appear from Table 2 to be the powerful influence upon syntactic selection which previous research had suggested it might be. Moreover, it appears to have the power to influence the level of linguistic performance of even linguistically 'deprived' groups. Previous research, which we reported earlier, had suggested that, although stylistic variation may appear in working-class language, it will generally be less extensive there than in middle-class language. Our findings, limited though they are, suggest that this generalization may not always hold true. A great deal depends upon the context in which language is produced. The implications of this go very deep, in fact, to the basic assumptions upon which the concept of two separable and class-related codes was originally raised.

The educational implications of our findings are more difficult to determine. Labov (1969) has stressed the complexity of much Negro and working-class language and has talked of its 'rich array of grammatical devices'. Even Bernstein has spoken on occasions of the concrete and metaphorical aesthetic of working-class language which gives it a formidable descriptive power. Seen in the light of these remarks our findings do not seem quite so surprising for they show us an increase in the grammatical complexity of working-class language when it is focussed upon a very limited and impersonal form of descriptive writing. In our corpus we have found many examples of highly complex nominal modification and qualification (e.g. '*Ther are 23 thin extensiv spoks on this wheel I describd*'; '*The oil which circulates around and around and lubricates is easy to obtan*') and of well-constructed passives (e.g. '*One the pedall as ben pushd and the weel as bene started moving*'). Apart from the spelling, there is little that is non-standard about these examples of grammatical complexity and, in fact, it is difficult to see them as, in any sense, an outgrowth of the aesthetic of the spoken dialect which these subjects used from day to day. It seems more likely that there we are dealing with the vernacular but with one of Gumperz's (1968) 'superposed varieties', a limited technical register which has been taught to the subjects and which they recognize as the norm for communication in particular situations. To go on from here to suggest that command over this technical register was given to our subjects in their schools or at their local technical college would, perhaps, be to overstate our case. On the other hand, it is worth remembering Robinson's (1965a) statement that, if working-class children are going to acquire a measure of elaboration in their language, they are more likely to do so in their writing than in their speech since it is only in the former that they are given much practice in their schools. This is not to say that Lawton (1968) was necessarily mistaken when he said that social class differences in written language will usually be greater than those found in speech. It implies, merely, that, even in working-class language which appears impoverished when judged by formal criteria, there may exist certain limited but very real areas of elaboration and complexity. It is into the analysis of such areas rather than into the construction of models to explain the linguistic 'deprivation' of working-class adolescents that we feel the energies of researchers need to be channelled.

References

BERNSTEIN, B. (1958). 'Some sociological determinants of perception: an enquiry into sub-cultural differences', *British Journal of Sociology* 9, 159–74.

BERNSTEIN, B. (1959). 'A public language: some sociological implications of a linguistic form', *British Journal of Sociology*, 10, 311–26.

BERNSTEIN, B. (1960). 'Language and social class', *British Journal of Sociology*, 11, 271–6.

BERNSTEIN, B. (1961). 'Social class and linguistic development: a theory of social learning', in HALSEY, A. H., FLOUD, J. and ANDERSON, C. A. (Eds.), *Education, Economy and Society*, 288–314. New York: Free Press.

BERNSTEIN, B. (1962a). 'Linguistic codes, hesitation phenomena and intelligence', *Language and Speech*, 5, 31–46.

BERNSTEIN, B. (1962b). 'Social class, linguistic codes and grammatical elements', *Language and Speech*, 5, 221–40.

BERNSTEIN, B. (1971). *Class, Codes and Control: Vol. 1: Theoretical Studies Towards a Sociology of Language*. London: Routledge and Kegan Paul.

COULTHARD, M. (1969). 'A discussion of restricted and elaborated codes', *Educ. Rev.*, 22, 1.

GUMPERZ, J. (1968). 'The Speech Community', in *International Encyclopaedia of the Social Sciences*, 381–6. London: Macmillan.

HALLIDAY, M. A. K. (1961). 'Categories of the Theory of Grammar', *Word*, 17, 241–92.

HALLIDAY, M. A. K., MCINTOSH, A. and STREVENS, P. (1964). *The Linguistic Sciences and Language Teaching*. London: Longmans.

HENDERSON, D. (1970). 'Social class differences in form-class usuage among five-year-old children', in BRANDIS, W. and HENDERSON, D. *Primary Socialization, Language and Education, Vol. 1: Social Class, Language and Communication*. London: Routledge and Kegan Paul.

LABOV, W. (1969). 'The logic of non-standard English', *Georgetown Monographs on Language and Linguistics*, 22, 1–22.

LAWTON, D. (1963). 'Social class differences in language development: a study of some samples of written work', *Language and Speech*, 6, III.

LAWTON, D. (1964). 'Social class language differences in group discussions', *Language and Speech*, 7, III.

LAWTON, D. (1965). 'Social class language differences in individual interviews', *Mimeograph, Sociological Research Unit, University of London Institute of Education.*

LAWTON, D. (1968). *Social Class, Language and Education*. London: Routledge and Kegan Paul.

LEECH, G. N. (1966). *English in Advertising: a Linguistic Study of Advertising in Great Britain.* London: Longmans.

QUIRK, R. (1962). *The Uses of English.* London: Longmans.

ROBINSON, W. P. (1965a). 'The elaborated code in working class language', *Language and Speech*, 8, 243.

ROBINSON, W. P. (1956b). 'Cloze procedure as a technique for the investigation of social class differences in language usage', *Language and Speech*, 10, 88.

SVARTVIK, J. (1966). *On Voice in the English Verb.* The Hague: Mouton.

VYGOTSKY, L. S. (1962). *Thought and Language.* London and Cambridge, Massachusetts: MIT Press.

SOCIAL SCIENCE RESEARCH COUNCIL GRANTS

The Social Science Research Council, through its Educational Research Board is prepared to consider applications for grants in support of research in all aspects of education. The Board is particularly keen to support projects which make a theoretical contribution as at least part of their results. It is not only large-scale projects involving research teams which are eligible for support: the Board is equally interested in considering applications for small projects with modest budgets. Under the provisions for SSRC personal research grants, it is possible for the salary of the principal investigator to be paid as well as some supporting costs. These are limited in number however, and research grants are chiefly intended to cover travel, subsistence, research assistance (including the appointments of full-time research fellows or assistants), necessary equipment and materials, secretarial help, etc.

Application for research grants of under £6,000 may be submitted at any time, while the closing dates for larger applications are: Grants between £6,000 and £50,000: 15 January, 15 April and 30 September; grants over £50,000; 31 August. Personal research grants: 30 September.

Intending applicants should first contact: Mrs A. V. S. Pope, Secretary to the Educational Research Board, SSRC, State House, High Holborn, London WC1R 4TH from whom full details and application forms may be obtained.

Social Class and Educational Research

T. S. Robertson formerly National Foundation for Educational Research

Summary
Social class has always been one of the most prominent variables in educational research, yet the various ways of classifying it have never ceased to be controversial. The author reviews the major classifications, assesses their validity, surveys the use made of them in research literature, suggests how they should be interpreted, and assesses to what degree findings couched in these terms correspond to the reality of social differences in Britain.

Social Stratification Theory

It has been repeatedly demonstrated that most people perceive themselves as belonging to a society in which there are different levels of prestige, and within which they themselves occupy a specific position (Kahan, Butler and Stokes, 1966; Martin, 1954; Goldthorpe and Lockwood, 1963). Awareness of social status has been shown to exist from an early age (Jahoda, 1959; Himmelweit, Halsey and Oppenheim, 1952). There has developed an extensive literature on this social stratification and some useful introductions to the topic are given by Jackson and Curtis (1968), Musgrove (1965, pp. 48-63), Corwin (1965, pp. 155-199) and Barber (1968). Bendix and Lipset (1965) include important contributions to social stratification theory by Sorokin 'What is a social class?', Parsons 'A revised analytical approach to the theory of social stratification' and Kornhauser 'The Warner approach to the theory of social stratification'. Frankenburg (1966) examines some alternative views of social status and himself adopts a Marxist view. There is a useful chapter on social stratification in Carr-Saunders, Jones and Moser (1958). Much of the work in this country stems from Stevenson (1928) whose contribution is referred to in the Registrar-General's Classification of Occupations (1966). Glass (1961) provides a historical account of stratification theory and Murphy (1964) reviewed developments from 1959 to 1964.

Two distinct strands can be perceived in research on stratification. One is the improvement of description and analysis of the phenomenon itself, the other is the study of various aspects of societies, such as health, education, sexual behaviour and political attitudes, comparing these for people of different social class. The former might be regarded as 'pure' research in stratification theory, the latter as 'applied'. The application of stratification theory in educational research forms the subject of this review.

There is now thorough documentation of many ways in which social status is important in studies of education, generally demonstrating the lower performance and opportunity of lower class children compared with higher class children. [For example Wall (1968, pp. 35-39); Neff (1938); Douglas (1964); Wiseman (1964, p. 38); Daniels (1964); Horobin, Oldman and Bytheway (1967); Wolfle (1961); Halsey (1961); Burt (1937) and Himmelweit (1966)].

While earlier attempts to dissect the ways in which these social class differences arose found relationships between low educational achievement and poverty, poor housing and malnutrition, more recent studies have concentrated on more subtle reasons for differences of attainment, such as attitudes, values and aspirations (Goodacre, 1968; Wall, 1968).

The Measurement of Social Class

Glass and Hall (Glass, 1954) refer to their 'belief, supported by other material contained in this volume, that in Great Britain occupation is probably the most important single criterion of status'. Kahan, Butler and Stokes (1966) in a study of social status and self appraisal by over

2,000 British electors found 'responders overwhelmingly assigned allocation to occupation (but with education, income, parentage and consumption pattern also involved)'. Ideally, then, any scale of social status should incorporate contributions from all of these factors, and perhaps others also.

Earlier studies were concerned with determining whether social status could be considered as comprising a continuous linear scale or whether there were natural discontinuities giving rise to a series of well defined levels, and, if so, how many of these levels or classes there were (see for example Rogoff, 1961; Goldthorpe and Lockwood, 1963; Sorokin, 1953). More recently it has been recognized that 'status' will have different meanings to respondents who may be asked to rate particular occupations. Perhaps there is no single concept of 'status'. Goldthorpe and Hope (1970) have carried out a pilot study rating separately, standard of living, prestige in the community, power and influence over others, and value to the society. They have found that these dimensions were differentially employed by the average rater, a result which throws doubt upon the value of some studies of 'occupational prestige'. Corwin (1965) has suggested a mosaic representation of social stratification to make allowance for the multidimensional nature of status. It is the author's opinion that for some research the concept of a linear scale is outdated, but that for many other studies it will continue to be used justifiably.

The Classification of Occupations

Since occupation is so important as a criterion in social class, much of social class classification depends on the adequate classification and rating of occupations. There is an international classification (International Labour Office, 1968), but this was formulated after most countries had already devised their own. In this country there are several basic classifications in which a status ordering is deliberately or implicitly present, and most studies use one or another of these, often modified in some way. They are:

Registrar-General's Social Class (General Register Office, 1966)

Registrar-General's Socio-Economic Group (ibid.)

Market Research Society (Market Research Society, 1963)

Hall-Jones (Hall and Moser, 1953).

Hall and Ziegel (1953) compared four national classifications (England and Wales, seven-fold; Italy, six-fold; France, seven-fold and United States, nine-fold). Another study involving international comparisons was that of Havighurst (1961) (U.S., Great Britain, Australia and Brazil). Stephenson (1958) used the Hall-Jones scale in a comparison of British and US data. The NFER has more recently been involved in the international evaluation of educational achievement where it has been necessary to devise an internationally acceptable scale (Pidgeon, 1967; Husén, 1967).

A common feature of the various classifications and derivations of them is that they are frequently revised. This is essential to reflect the changing nature of employment, but often makes comparison difficult. For example, the occupational classification adopted for the Crowther Report (CACE, 1959) was based upon the Registrar-General's 1951 socio-economic group. This was revised in 1961, and later studies based upon socio-economic group may not be exactly comparable.

British Classifications

The most widely used classification in Britain is the Registrar-General's social class. It has the following five categories:

I Professional, etc.
II Intermediate
III Skilled
IV Partly skilled
V Unskilled

Despite widespread justified criticism, it is still much used. To quote from the Registrar-General (1966) . . . 'As far as possible, each category is homogeneous in relation to the basic criterion of the general standing within the community of the occupations concerned . . . correlated with other factors such as education and economic environment . . . no direct relation to average level of remuneration'. 'The present social class grouping follows the same general lines adopted in 1921, 1931 and 1951, although the allocation of occupations to five groups has varied from census to census in accordance with changes in economic conditions and with the intention of preserving the gradient rather than literal continuity.'

Criticism of the Registrar-General's social class centres upon the heterogeneity of some classes, and anomalies such as the grouping of all farmers in II, irrespective of the size of their holdings. All school teachers are in II (intermediate) when I (professional) might seem more appropriate. Sales representatives, many of them highly qualified, are all in III (skilled). The inclusion in the more recent censuses of questions about the size of businesses has improved the data for many employers and farmers, and the introduction of the new classification by socio-economic group takes advantage of this improvement. However, once allocations have been made, there is no way of converting socio-economic classification to social class or vice-versa, unless by reference back to the original data. The sub-division of social class into manual and non-manual categories, undertaken by many workers (for example Kellmer Pringle *et al.*, 1966; Morton-Williams, 1967) and now facilitated by a division of this axis by the Registrar-General, removes some of the disadvantage of the social class classification. This converts the five-fold classification into a six-fold one, which is then frequently recombined to two-fold:

I Professional II Intermediate III Skilled non-manual	Middle Class
III Skilled manual IV Partly skilled V Unskilled	Working Class

It has often been claimed that a two-fold division is the most fundamental but the line of demarcation is variously placed by different workers (see later).

In favour of the use of social class is the comparative ease and precision of collecting and coding data using the Registrar-General's *Classification of Occupations*; and the large number of studies over many years in which it has been used, permitting comparability. In at least two studies (Goodacre, 1968; Horobin *et al.*, 1967) data have been classified both on Registrar-General's and other more refined classifications. When, in both cases, no appreciable difference was found, the Registrar-General's was adopted.

In 1951, the Registrar-General introduced a new classification, socio-economic group, utilizing the fuller data now being collected at the census. Groups include 'people whose social, cultural and recreational standards and behaviour are similar'. This 17-fold classification is too clumsy for most purposes and on the Ward Library Sheets (see later) a six-fold classification derived from it is used:

Professional (S-e G. 3 & 4)
Employers and Managers (including farmers) (S-e G. 1, 2 & 13)
Foremen, skilled manual and small farmers (S-e G. 8, 9, 12 & 14)
Intermediate and junior non-manual (S-e G. 5 & 6)
Personal service, semi-skilled and agricultural (S-e G. 7, 10 & 15)
Unskilled (S-e G. 11)

A difficulty of using socio-economic-group-based classifications in educational research is the need to obtain precise information about numbers of employees, in the case of some employers. Practical problems of this kind are to be discussed later in this paper.

In addition to the Registrar-General's two classifications there is one other scheme that has been widely used for educational research in this country. This is the Hall-Jones scale (Hall and Jones, 1950; Hall and Moser, 1953). It improves the discrimination in the non-manual areas, but leaves the three manual classes (skilled, semi-skilled and unskilled) intact to provide a seven-fold scale. It has the advantage of having been tested against the actual attitudes of a large number of responders whose assessments of the social status of a list of occupations was determined. The descriptions of the classes are:

1. Professional and high administrative
2. Managerial and executive
3. Inspectional and supervisory, higher grade
4. Inspectional and supervisory, lower grade
5. Skilled manual and other grades of non-manual
6. Semi-skilled manual
7. Unskilled manual

Various modifications of the Hall-Jones scale have been used. For example, Martin (1954) split off the weekly paid non-manual jobs in 5 from the skilled manual and called these 4(a). He asked his subjects to give their own estimate of which class they occupied and formed the following scale based on self-subjective status.

Martin	*Hall-Jones Modified*	*Subjective Status of Respondents*
A	1 and 2	Middle class
B	3, 4 and 4(a)	Middle class
C	3, 4 and 4(a)	Working class
D	5, 6 and 7	Middle class
E	5, 6 and 7	Working class

Dockrell (1966) combined A, B and C of these as 'middle class' and D and E as working class, despite the subjective status element. Himmelweit (1954) derived the following:

Upper Middle Class (Hall-Jones 1)
Middle Middle Class (Hall-Jones 2 and 3)
Lower Middle Class
Upper and Lower Manual Class

Little and Westergaard (1964) used the following grouping to compare results using the Hall-Jones classification with others using the Registrar-General's social class:

Hall-Jones	*Registrar-General*
1, 2, 3 and 4	I and II
5	III
6 and 7	IV and V

They also grouped Hall-Jones differently to compare with data in the Crowther Report (CACE, 1959):

Hall-Jones	*Crowther*
1–3	Professional/managerial
4–5	Other non-manual and skilled manual
6–7	Semi-skilled and unskilled

Stephenson (1958) used the Hall-Jones scale in a comparative study between America and Britain. Goodacre (1968) found for a rather special population that Hall-Jones and Registrar-General's Social Class correlated 0·77. Bechhofer (1969) draws a comparison between the three British classifications so far considered. The Hall-Jones scale has been used in several scales involving not only occupation but parental education and other characteristics, which will be discussed more fully later (Brandis and Henderson, 1970; Goodacre, 1968; Morris, 1966).

A number of researchers have referred to the 'Government Social Survey Classification' but this has not been traced by the author (Kahan, Butler and Stokes, 1966; Floud and Halsey, 1961; Floud, Halsey and Martin, 1956). Inquiry at the Social Survey Division, Office of Population Censuses and Surveys (formerly the Government Social Survey) brought the reply that various classifications had been used by this organization for different purposes.

Other British occupational classifications include Gooch and Kellmer Pringle (1966) in which four judges independently allocated jobs to four classes (upper and lower non-manual, upper and lower manual); Kahan, Butler and Stokes (1966) based on the Market Research Classification; Barker-Lunn (1970); Lindsay (n.d.).

A national report (Pidgeon, 1967) for the International Evaluation of Achievement used the nine-fold classification of IEA In an international report on this project (Husén, 1967 p. 144), Anderson lists these groups with examples, discusses the problems of international comparisons and describes a measure of *spread* of parental occupation of pupils on the school roll by calculating the standard deviation.

The Social Class Dichotomy

A dichotomous view of social class has repeatedly been shown to be fundamental (Kahan, Butler and Stokes, 1966; Goldthorpe and Lockwood, 1963; Douglas, 1964; Ford, 1969; Brandis and Henderson, 1970). A two-way split of social class (middle-class/working class or non-manual/manual) has most frequently been used in educational research —often by condensation of five-fold, seven-fold or nine-fold initial scales, and it is the position of the line of demarcation which is sometimes in dispute. Bernstein and Young (1967) used Brandis' (1966) Index of Social Class to separate five social classes 'two clearly middle class, two clearly working class and one marginal', then compared 351 families in two areas (mainly middle class and mainly working class). Kahan, Butler and Stokes (1966), using responders' self-appraisal, found the cleavage not between manual and non-manual but between skilled and supervisory non-manual and other (lower) non-manual. Martin (1954) found a considerable number of workers in the lower classifications of Hall-Jones scale assessed themselves as belonging to the middle class. Goldthorpe and Lockwood (1963), however, criticise self-appraisal by poll-type interview. Bernstein (1960 and 1965) has compared the restricted use of language in working classes compared with middle classes and Goldthorpe and Lockwood (1963) contrast working and middle class perspectives under the headings of 'general beliefs', 'general values' and 'attitudes on more specific issues'. The middle class tends to adopt a hierarchical view of society as having a number of relatively 'open' strata, whilst the working class tends to view society as being sharply divided into two contending sections.

It is interesting to note that any 'upper class' seems to be eliminated from or ignored in most social research, at least in this country. An amusing pastime of labelling social customs and language usage as 'U' and 'non-U' was in vogue some years ago and seems to have represented a dichotomy between 'upper class' and 'the rest' (see, for example, Mitford, 1949).

Social Status Scales Based on More than One Criterion

Those scales mentioned so far have been scales of occupational status. As already stated, the concept of status incorporates also elements of education, income, parentage and consumption pattern, and a number of scales have been devised both at home and abroad, to take all these contributions into account. Warner *et al.*, (1949) in America describe how to compute an index of status characteristics including weighting of scores on separate scales.

Duncan (1961) describes an index for all occupations. Baker, Schutz and Hinze (1961) in America used Warner's index aided by the much fuller data on pupils' cumulative records than is usually available in this country. Also in America, Sewall and Shah (1967) devised a weighted combination of occupation, father's and mother's occupation and education, income and wealth and briefly described their statistical method, and Rogoff (1961) constructed a set of ranked classes based on father's occupation and education, and siblings' occupation.

Brandis (1966) in Britain has developed two new indices of social class, based on mother's and father's occupation and education, important in that careful consideration was given to their mathematical derivation. These were used in a study of social class, language and communication by Brandis and Henderson (1970, pp. 130-136).

Douglas (1964) discusses the shortcomings of classification solely by occupation, and later (Douglas, Ross and Simpson, 1968) used a scale based on parental occupation and education. However the educational criterion was 'parent went to secondary school'. Since the term 'secondary school' has altered in meaning, and now a large number of pupils attend comprehensive schools, this definition would need revision if the scale were to be used again.

Other Scales involving Occupational Criteria

Parental occupation has been used as a variable in factor analysis of more general environmental features such as 'home circumstances'. Morris (1966) used father's and mother's occupation, cultural standard, stability, size of family, parental help and encouragement, leisure reading and other out-of-school activities as home circumstance factors. Young and McGeeney (1968, pp. 142-143) mention a regression analysis to find the contributions to a 'family background' variable of parental occupation, age of leaving school and further education. Kemp (1955) used such additional factors as learning musical instruments, possession of cars and telephones, along with parental occupation and education. The IEA used a variety of measures (Pidgeon, 1967; Husén, 1967).

School Neighbourhood Assessments

In many studies, assessments have been made of school neighbourhood. Whilst occupation pattern is an important feature of school neighbourhood, it is not the only one. Burt's (1937) classic study of the distribution of backwardness in London showed its relation to neighbourhood and poverty, and to material, intellectual, emotional and moral conditions of the home. Rogoff (1961) considered the minority children in neighbourhoods with particular social class pattern, and the effects on social mobility. Wiseman (1964) used a variety of neighbourhood characteristics, including medical and welfare records. Warburton (1964) had ratings made by school inspectors, etc., taking into account such factors as type and age of houses, and presence or absence of front gardens.

At one time the 'Juror-Index' (Gray, Corlett and Jones, 1951) was a useful neighbourhood measure, but changes in the liability for jury service have now invalidated it. Goodacre (1968) used a simple measure based on census data of social class which worked well in her limited study in London. Ross, Bunton, Evison and Robertson (1972) found a high degree of correlation between neighbourhood measurements based on social class, socio-economic group and a home-background measure devised by Sugarman. The Ward Library Sheets (available from the Census Customer Service, Fareham, Hants.) include a variety of measures (including social class and socio-economic group) which could, if desired, be used as neighbourhood measures.

Education and Some Social Class Differences

The amount of information now collected under this general description is too great to list in detail. Summaries are available in Wall (1968, pp. 35-39), Musgrove (1965, pp. 64-82), Wiseman (1964, p. 38) and Cullen (1969).

The Collection of Information on Social Class in Schools

It is unfortunate that this is one of the areas of inquiry where head teachers and local authorities are most sensitive. A flat refusal to allow children to be questioned about what their parents do for a living is sometimes met (Pidgeon, 1967, p. 313). The reasons are sound enough—the very children one would like to question are those where information is lacking from school records, usually because the father is dead, has left home or is in prison. This sensitivity must be foreseen in the planning of any research project. Nevertheless, high levels of co-operation, and sound data, can be secured by good fieldwork (for some general remarks on this topic see Robertson, 1968).

Common methods of collecting data include (1) access to pupils' record cards; (2) inclusion of questions on pupil-completed, teacher-completed or parent-completed questionnaires; (3) questions to teachers or parents in structured interviews.

In its simplest form an estimate of school social class may be obtained by a question such as that asked in Monks (1968): 'Please estimate the percentage of children in the September 1965 intake coming from homes in the following categories. . .'. There followed a grid and a sample list of occupations. The results may be crude but will be sufficient to perform an initial sorting process into schools with virtually no middle-class pupils and schools with some or many. In a later stage of the same project, school staffs were asked to supply a copy of the school roll with the name, age, address, sex, form, house, mother tongue, ethnic origin, ability rating and parental occupation of each child. Guidance was given as to the definitions of some of these. Although in some cases replies were not as complete as might be desired (Monks, 1970, p. 24), at a third stage of the project still closer questioning filled many gaps (Ross, Bunton, Evison and Robertson, 1972). The percentages of classifiable replies on parental occupation were as low as 20 in some schools and as high as 95 in others. In a sample of 363 pupils with initially 62 per cent useable replies, the third stage questioning increased the useable replies to 86 per cent. Examination of the unclassifiable cases showed the following categories:

1. No father (deceased, parents separated, orphans, etc.)
2. Information insufficient (e.g. 'works at AEI')
3. Unemployed
4. No information

In the Plowden Survey (Morton-Williams, 1967), if there was no father the occupation of the father or father-substitute when last in the household was used. In this way, in the favourable situation of face-to-face interview with the parent, only 1 per cent of families could not be classified. (There were in fact 6 per cent of cases where the natural father was absent.)

In the NFER French Project (Burstall, 1970), if there was no father the mother's occupation was used, but there are disadvantages in this, as a working mother may have to take employment well below her portentiality. In nearly all cases where the parental occupation had been left blank, further questioning revealed that the father had left home and it was possible to find his last occupation. If the child was 'in care' it was classified in the lowest social class. In this way, a high percentage of replies was classified. The difficulties of the collection of social class data are illustrated in one school in the NFER Comprehensive Project (Ross, Bunton, Evison and Robertson, 1972) in which secondary pupils of comparatively low ability were asked to pencil their address and father's occupation on the back of a questionnaire they were completing. There were 18 cases out of 72 where further information had to be sought. The reasons why the data were inadequate were:

No information given	6
Name of employer only	4
Description inadequate	6
No father	2

Further inquiry from the school gave three occupations and one 'no father', so the final percentage useable was 83.

In another project where a fairly superficial measure of school social class was needed, teachers were asked to place each child in one of five described social classes 'using such information as was available to the teacher', but this method cannot be recommended if precision is desired.

Colfox and Allen (1967) compared pre-coded and open-ended items in children's reports of their fathers' occupations (in America) and found open-ended preferable.

The Need for Comparability

Stacey (1969) covers the field of comparability in social research, including a chapter by Bechhofer on occupations. The need for comparability is perhaps obvious. Any piece of research should be designed so that it could be repeated by another competent worker to give the same result. It is also of little value if it produces some isolated fragment of information which, because not comparable, cannot be related to any other. Comparability in research on social class implies using an existing classification, or devising a classification which later workers can use. Existing classifications are unsatisfactory in a number of ways. They attempt to represent social status as a linear scale, when it is in fact multi-dimensional. They are largely forced by practicalities into using one or two of the prominent factors (occupation and education) contributing to status. Major occupational groups are too heterogeneous, and probably too much effort has gone into subdivision of the upper than the lower social classes, perhaps because social scientists are not familiar enough with the differences of culture between the many trades collectively called 'skilled', 'semi-skilled' and 'unskilled'. (For an example of an interesting study of one trade, the compositor's, see Cannon, 1967.) Fraser (1959) devised a useful eight-fold classification that improved the separation at the lower end and which could be useful in future studies. The occupational structure varies between regions and countries, and is said to be undergoing major changes in this country at the moment, so geographical and temporal comparability are difficult to achieve. For example Douglas, Ross and Simpson (1968) found in a longitudinal study that 40 per cent of the fathers of pupils in their sample had changed their social class in 11 years. When educational criteria are to be used it is difficult to achieve comparability, as school organization changes and the statutory minimum leaving age rises. Measures based on type of school attended, and age of leaving school, are valuable but become outdated.

Discussion

This paper has reviewed briefly the theory of social stratification and mentioned the inadequacies of a linear scale of social status. It has emphasized the dichotomous view of social status into 'middle' and 'lower' classes. Research is moving from recording the differences of opportunity and achievement of pupils from different levels towards analysis of why these differences arise or exist, in terms of differences in culture, attitude and environment operating from an early age.

Increasing understanding of why attainment differs should lead to action aimed at reducing the difference by compensatory education. It seems that measures such as movement towards nursery education and comprehensive schools could go some way, but not all the way, towards redressing the balance.

In educational research, a variety of measures of social status have been used, all with a strong occupational content. Social status variables have been incorporated in factor analyses to detect more general variables such as 'home circumstances' and 'family background'. In all social status studies, compromise has to be made between what is desirable in terms of accuracy and detail and what can reasonably be requested and obtained. Because detailed and exact instructions for classification exist, the Registrar-General's classification of occupations is probably the best for general use in educational research in Britain. It has the advantage of generating results that are comparable with those of many other studies. For many purposes the somewhat dated social class grouping continues to be of great use, particularly if the skilled workers in Social Class III are subdivided into manual and non-manual, to give a six-fold scale, recombinable to two for some statistical purposes. In rural schools the placing of all farmers in Social Class II may be misleading, in public schools there may be insufficient separation at the top and in working-class schools insufficient separation at the lower half of the scale. Use of the Registrar-General's socio-economic group, or of a scale such as that of Fraser (1959) might be desirable in these cases.

References

BAKER, R. L., SCHUTZ, R. E. and HINZE, R. A. (1962). 'The influence of mental ability on achievement when socio-economic status is controlled', *J. Exper. Educ.*, **30**, 255–8.

BARBER, B. (1968). 'Social stratification. Introduction', in SILLS, D. L. (Ed.) *International Encyclopaedia of the Social Sciences*. Vol. 15. New York: MacMillan and the Free Press.

BARKER LUNN, J. C. (1970). *Streaming in the Primary School*. Slough: NFER.

BECHHOFER, F. (1969). 'Occupations', in STACEY, M. (Ed.) *Comparability in Social Research*. London: Heinemann.

Bendix, R. and Lipset, S. M. (Eds.) (1965). *Class, Status and Power*. New York: Free Press.

Bernstein, B. (1960). 'Language and social class', *Brit. J. Sociol.*, **11**, 271–6.

Bernstein, B. (1965). 'A socio-linguistic approach to social learning', in Gould, J. (Ed.) *Penguin Survey of the Social Sciences*. London: Penguin.

Bernstein, B. and Young, D. (1967). 'Social class differences in conceptions of the use of toys', *Sociology*, **1**, 2.

Brandis, W. (1966). 'Indices of Social Class, Communication and Control: Research Memoranda'. Sociological Research Unit, London University Institute of Education.

Brandis, W. and Henderson, D. (1970). *Social Class, Language and Communication*. London: Routledge and Kegan Paul.

Burstall, C. (1970). *French in the Primary School*. Slough: NFER.

Burt, C. (1937). *The Backward Child*. London: University of London Press.

Cannon, I. C. (1967). 'Ideology and occupational community: a study of compositors', *Sociology*, **1**, 2, 165–85.

Carr-Saunders, A. M., Jones, D. C. and Moser, C. A. (1958). *Social Conditions in England and Wales*. Oxford: Clarendon Press.

Central Advisory Council for Education (1959). *15 to 18* (Crowther Report). London: HMSO.

Colfox, J. D. and Allen, I. L. (1967). 'Pre-coded versus open-ended items and children's reports of father's education', *Sociol. Educ.*, **40**, 1, 96–8.

Corwin, R. G. (1965). *A Sociology of Education*. New York: Meredith.

Cullen, K. (1969). *School and Family*. Dublin: Gill and MacMillan.

Daniels, J. C. (1964). 'Social class and verbal intelligence', *New Society*, **21**:v:64, 8.

Department of Education and Science: Central Advisory Council for Education (England) (1967). *Children and their Primary Schools* (Plowden Report). London: HMSO.

Dockrell, W. (1966). 'Secondary education, social class and the development of abilities'. *Brit. J. Educ. Psychol.*, 36, 7–14.

Douglas, J. W. B. (1964). *The Home and the School*. London: MacGibbon and Kee.

Douglas, J. W. B., Ross, J. M. and Simpson, H. R. (1968). *All Our Future*. London: Peter Davies.

Duncan, O. D. (1961). 'A socio-economic index for all occupations', in Reiss, A. *et al.*, *Occupation and Social Status*. New York: Free Press.

Floud, J. and Halsey, A. H. (1961). 'Social class intelligence tests and selection for secondary schools', in Halsey, A. H., Floud, J. and Anderson, C. E. (Eds.) *Education, Economy and Society*. New York: Free Press.

Floud, J. E., Halsey, A. H. and Martin, F. M. (1956). *Social Class and Educational Opportunity*. London: Heinemann.

Ford, J. (1969). *Social Class and the Comprehensive School*. London: Routledge and Kegan Paul.

Frankenburg, R. (1966). *Communities in Britain*. London: Penguin.

Fraser, E. D. (1959). *Home Environment and the School*. London: University of London Press.

General Register Office (1961). *Occupation Tables*. London: HMSO.

General Register Office (1966). *Classification of Occupations*. London: HMSO.

Glass, D. B. (Ed.) (1954). *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Glass, D. V. (1961). 'Education and social change in modern England,' in Halsey, A. H., Floud, J. and Anderson, C. A. (Eds.) *Education, Economy and Society*. New York: Free Press.

Goldthorpe, J. H. and Hope, K. (1970). 'Occupational prestige', in *Synopses of papers prepared for the Nuffield Social Mobility Project, 1968–70 (Occupational Change in a Generation)*. Oxford: Nuffield College (duplicated sheets).

Goldthorpe, J. H. and Lockwood, D. (1963). 'Affluence and the British class structure', *Sociol. Rev.*, **9**, 133–64.

Gooch, S. and Kellmer Pringle, M. L. (1966). *Four Years On*. London: Longmans.

Goodacre, E. J. (1968). *Teachers and their Pupils' Home Background*. Slough: NFER.

Gray, P. G., Corlett, T. and Jones, P. (1951). *The Proportion of Jurors as an Index of the Economic Status of a District*. London: Central Office of Information.

Hall, J. and Caradog-Jones, D. (1950). 'The social grading of occupations', *Brit. J. Sociol.*, **1**, 1.

Hall, J. R. and Moser, C. A. (1953). 'The social grading of occupations', in Glass, D. (Ed.) *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Hall, J. R. and Ziegel, W. (1953). 'Comparison of social data for England and Wales, Italy, France and the USA', in Glass, D. (Ed.) *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Halsey, A. H. (1961). *Ability and Educational Opportunity*. Paris: OECD Publications.

Havighurst, R. J. (1961). 'Education and social mobility in four societies', in Halsey, A. H., Floud, J. and Anderson, C. A. (Eds.) *Education, Economy and Society*. New York: Free Press.

Himmelweit, H. T. (1954). 'Social status and secondary education since the 1944 Act', in Glass, D. (Ed.) *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Himmelweit, H. T., Halsey, A. H. and Oppenheim, A. N. (1952). 'The views of adolescents on some aspects of the social class structure', *Brit. J. Sociol.*, **3**, 148–72.

Horobin, G., Oldman, D. and Bytheway, B. (1967). 'The social differentiation of ability', *Sociology*, **1**, 2, 113–29.

Husen, T. (Ed.) (1967). *International Achievement in Mathematics*, Vols. I and II. New York: Wiley.

International Labour Office (1968 revised 1969). *International Standard Classification of Occupations*. Geneva: ILO.

Jackson, E. F. and Curtis, R. F. (1968), in Blalock, H. M. and Blalock, A. B. *Methodology in Social Research*. New York: McGraw Hill.

Jahoda, G. (1959). 'Development of the perception of social differences in children from six to ten', *Brit. J. Psychol.*, May, **50**, 159–75.

Jones, D. C. (Ed.) (1934). *The Merseyside Social Survey* (3 vols.). Liverpool University Press. Hodder and Stoughton.

Kahan, M., Butler, D. and Stokes, D. (1966). 'On the analytical division of social class', *Brit. J. Sociol.*, **17**, 2, 124.

Kellmer Pringle, M. L., Butler, N. R. and Davie, R. (1966). *11,000 Seven-Year-Olds*. London: Longmans.

Kemp, L. C. D. (1955). 'Environmental and other characteristics determining attainments in primary schools', *Brit. J. Educ. Psychol.*, **25**, 67–77.

Kornhauser, R. R. (1965). 'The Warner approach to social stratification', in Bendix, R. and Lipset, S. M. *Class, Status and Power*. New York: Free Press.

Lindsay, C. (n.d.). *School and Community*. Oxford: Pergamon.

Little, A. and Westergaard, J. (1964). 'The trend of class differentials in educational opportunity in England and Wales', *Brit. J. Sociol.*, **15**, 4, 301–16.

Market Research Society (1963). *The Social Class Definition in Social Research*. London: The Market Research Society.

Martin, F. M. (1954). 'Some subjective aspects of social stratification', in Glass, D. V. (Ed.) *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Mitford, N. (1949). *Pursuit of Love*. London: Hamilton.

Monks, T. G. (1968). *Comprehensive Education in England and Wales*. Slough: NFER.

Monks, T. G. (Ed.) (1970). *Comprehensive Education in Action*. Slough: NFER.

Morris, J. M. (1966). *Standards and Progress in Reading*. Slough: NFER.

Morton-Williams, R. (1967). 'The 1964 national survey: survey among parents of primary school children', in Department of Education and Science: Central Advisory Council for Education (England). *Children and their Primary Schools*. Vol. 2 (Plowden Report). London: HMSO.

Murphy, R. J. (1964). 'Some recent trends in stratification theory and research', *Annals of the American Academy of Political and Social Science*, **356**, November Supplement, 142–67.

Musgrove, P. W. (1965). *The Sociology of Education*. London: Methuen.

Neff, W. S. (1938). 'Socio-economic status and intelligence: a critical survey', *Psychol. Bull.*, **35**, 727–57.

Parsons, T. (1952). *The Social System*. Glencoe, Ill.: Free Press.

Parsons, T. (1965). 'A revised approach to the theory of social stratification', in Bendix, R. and Lipset, S. M. (Eds.) *Class, Status and Power*. New York: Free Press.

Pidgeon, D. (1967). *Achievement in Mathematics*. Slough: NFER.

Robertson, T. S. (1968). 'Some practical considerations in the conduct of educational research in the United Kingdom', *Statistician*, **18**, 2, 129–40.

Rogoff, N. (1961). 'Local social structure and educational selection', in Halsey, A. H., Floud, J. and Anderson, C. A. (Eds.) *Education, Economy and Society*. New York: Free Press.

Ross, J. M., Bunton, W. J., Evison, P. and Robertson, T. S. (1972). *A Critical Appraisal of Comprehensive Education*. Slough: NFER.

Sewell, W. H. and Shah, V. P. (1967). 'Socioeconomic status, intelligence and the attainment of higher education', *Sociol. Educ.*, **40**, 1, 1–23.

Sorokin, P. A. (1965). 'What is a social class?', in Bendix, R. and Lipset, S. M. (Eds.) *Class, Status and Power*. New York: Free Press.

Stacey, M. (Ed.) (1969). *Comparability in Social Research*. London: Heinemann.

Stephenson, R. M. (1958). 'Stratification, education and occupational orientation, a parallel study', *Brit. J. Sociol.*, **9**, 42–52.

Stevenson, T. H. C. (1928). 'The vital statistics of wealth and poverty', *J. Roy. Statis. Soc.*, Series A, **91**, 207–30.

Wall, W. D. (1968). *Adolescents in School and Society*. Slough: NFER.

Warburton, F. W. (1964). 'Attainment and the school environment', in Wiseman, S. *Education and Environment*. Manchester University Press.

Warner, W. L. *et al.* (1949). *Social Class in America*. New York: Harper and Row.

Young, M. and McGeeney, P. (1968). *Learning Begins at Home*. London: Routledge and Kegan Paul.

Wolfle, D. (1961). 'Educational opportunity, measured intelligence and social background', in Halsey, A. H., Floud, J. and Anderson, C. A. (Eds.) *Education, Economy and Society*. New York: Free Press.

Wiseman, S. (1964). *Education and Environment*. Manchester University Press.

Educational Research in Britain 3

Edited by H. J. Butcher and H. B. Pont

This is the third volume discussing current educational research in Britain. Work of research centres and projects – such as The Humanities Curriculum Project – is described: other topics range from child-rearing practices to science teaching: from further education to personality and academic attainment.

£4.60 net **0 340 16792 0**

Available through all good bookshops

UNIVERSITY OF LONDON PRESS LTD

St Paul's House Warwick Lane London
EC4P 4AH

Objectives in the Humanities

Gabriel Chanan, National Foundation for Educational Research

Summary
A key element in curriculum theory over recent decades has been a particular notion of teaching objectives as specifiable intentions on the part of the teacher. It is argued that this notion fails to take account of objectives generated in the course of teaching. These are particularly important in some humanities subjects, and perhaps their neglect is attributable to the tendency of curriculum theory to take its primary models from the sciences. This article seeks to clear a way for the discovery of a mode of describing objectives which would be more suitable to humanities subjects. Some weak points in previous attempts to do this are analysed, and some of the major problems of the field are presented.

Limitations of the linear model

The notion of objectives which has been prevalent in educational theory and research over the last twenty years has at root been that of one phase in a simple, linear, purposive activity on the part of the teacher. His aim for the pupil is such and such, his methods of attempting to achieve it are such and such, the results are such and such. If teachers did not proceed in this manner, their classroom activities must be governed uncritically by tradition, intuition or inertia. It has seemed, to many theorists, no more than logical to recommend that this pattern of procedure, analogous to that followed in most educational research, should be instituted as standard classroom practice: the formal stating of goals (in the light of pupils' present level of attainment, objectively assessed), the selection of methods, and the evaluation of results. This recommendation is now proselytized in a number of books aimed at the practising teacher.[1] However, like research studies themselves, it meets with a certain amount of scepticism from teachers and educationists.

It would be a mistake to attribute this scepticism solely to intellectual inertia. The individual teacher is often not in a *position* to implement advocated reforms even if he wants to: the harried circumstances of daily school life frequently do not permit the carrying-through of medium or long-term orderly procedures.[2] However, even if time and conditions in the average school were ideal, it is questionable whether our current image of the function of educational objectives would be found suitable. It is far too simple for any purposeful human activity, let alone the particularly complex one of teaching.

Objectives are generated not merely as intentions prior to the act of teaching: they are generated by the dawning realization of certain needs and opportunities *during* the act of teaching. The prevalent notion of objectives recognizes classroom 'feedback' only as a source of information on which to base the next formal statement of objectives. What tends to be neglected is the fact that in the light of this recurrent on-the-spot feedback the teacher is continually modifying his goals and methods in the act of teaching itself, so that the original statement of objectives may be virtually obsolete by the end of a single lesson, let along a week or term of lessons.

Differences between sciences and humanities

At this point it becomes necessary to distinguish between the sciences and the humanities. Flexibility of teaching goals is more essential to the teaching of humanities subjects (including the social sciences), more intrinsic to it, than to the teaching of mathematics and the pure sciences, however much the latter subjects would also benefit by it. Curriculum evaluation methods have taken their cue primarily from mathematics and science subjects precisely because it appeared to be easier to structure these subjects into a hierarchy of specific goals. Integral to this approach was the assumption that the goals were implicit in the subject matter. The objectives would amount to mastery of the particular concepts of which that subject matter was composed: 'The essence of this approach . . . is the placing of each pupil, individually, in an ordered sequence of situations which will enable him to learn most effectively. . . . This presupposes that it is possible, in most subjects, to postulate at least the basis of an order of conceptual learning, such that at any point in the learning sequence no major principle or concept is presented for learning which subsumes other concepts not already understood.'[3]

For humanities subjects, however, it is by no means a foregone conclusion that *an* order of conceptual learning is to be found; it is equally uncertain, therefore, that an 'ordered sequence of situations' can be determined. Indeed, if one such was determined, it might represent a step backwards, for what characterizes the subject matter of the humanities is specifically the *multiplicity* of paths by which it may be approached. Humanities teachers have therefore concentrated, on the one hand, on stimulating *any* form of involvement with suitable subject matter, and on the other on alerting themselves to recognize *any* of a broad range of desirable learning outcomes, including outcomes that they could not have predicted.

It is not, therefore, surprising if, on being asked their objectives, humanities teachers tend to produce only an assembly of vague aims. While this may conceal intellectual inertia, it may equally well represent an awareness that intentional aims before the act of teaching do not necessarily amount to much more than opening gambits, compared with the rapid crystallization of aims that occurs again and again, and differently each time, during the classroom encounter itself.

To what extent this description of teaching could apply just as well to science subjects is something on which those more intimate with those fields might care to pronounce. Criticism of the classic curriculum evaluation approach has certainly been heard from science-teaching quarters.[4]

In research circles various alternatives to the 'input-outcome' model of teaching have been advanced, such as an 'anthropological' model,[5] action research[6] and participant-observer research[7]. It is not possible here to evaluate the varied achievements of these schools of thought. It is sufficient for the present discussion to observe, firstly, that the question of objectives is still very much an open one and, secondly, that the classical 'input-outcome' model of teaching is still predominant in published educational research and also in the field of 'educational technology' which has been influenced by it.[8]

Educational research and technology alike claim scientific status, and it is difficult for the humanities teacher to judge whether the inappropriateness for his subject which he discerns in their approach should be attributed to their origins in scientific thinking as such or to their particular interpretation of scientific thinking. In either event, he is likely to have noticed that many of the key words in the classic curriculum evaluation approach carry a meaning significantly different from that which they carry in his own context: 'evaluation' is used to mean *measurement*, whereas to him it means *judgment by reference to social, moral or aesthetic values*—a distinctly non-measuring activity; 'affective' objectives are frequently used by the curriculum theorists to mean no more than objectives concerning the pupils' feelings towards the cognitive subject matter, whereas to the humanities teacher the term would also mean objectives concerning the pupils' emotions in themselves, i.e. in relation to the experience of life in general. Again, the humanities teacher might encounter the emphasis on individualized learning which appears in the quotation above (note 3) and regard this as a threat to the reciprocity of learning which he hopes to facilitate by means of classroom discussion. (He would share, however, the implied aversion to the imposed *uniformity* of learning of the traditional classroom.) Principally, he would be unlikely to share the frequent assumption that the objectives of learning are unambiguously inherent in the content of the given subject matter.

The content of history, literature, philosophy, religion and social studies is infinite and open-ended, and any one item of study in these areas can be approached at many levels, from the simplest to the most complex. The precise goals of teaching a certain subject matter to a certain class are to be found in a selected treatment of the subject matter, not simply in the content of the subject matter 'itself'. One might read *Animal Farm*, learn about Cromwell or discuss civil rights with the first form as well as the sixth, having significantly different (though ultimately related) goals in view. Equally important, it is only *by* engaging in any of these activities with a particular class that one would expect to be able to discover just what educational value the activity could have for that class.

This does not make evaluation superfluous. What it does mean is that the teacher's ability to recognise when any of a number of possible beneficial results is emerging during the classroom encounter (or in written work) is more important than his having been able to specify some of them in advance. The exercise of specifying goals in advance may well have been beneficial in clarifying his thinking and his inferences from preceding experience, and in rationalizing his teaching initiatives. But it would have become an impediment if it had been taken to define the function and scope of a particular course, for it would then have impaired his capacity to recognize and exploit fruitful deviations arising from pupils' reactions.[9]

Unpredictable objectives

The capacity to recognize unplanned beneficial outcomes during their very emergence is neither marginal to good teaching nor purely a matter that must be left to intuition. An important part of the art of teaching may consist in capitalizing upon these unpredictable opportunities, and there is no reason to suppose that this capacity could not be increased with training. Many other activities of life are at the same time purposive and yet open-ended as to their beneficial outcomes. To go to the cinema or theatre, to meet friends, to read a book or to go on a holiday is to undertake an activity in expectation of pleasure or reward when it could clearly be impossible or self-obstructive to lay down in advance the exact nature of the pleasure or reward to be obtained. Yet equally clearly the faculty in us which recognizes the pleasure or reward when it appears, and which thereby permits us to exploit it far more fully, is susceptible of deliberate amplification and refinement.

The objectives which the humanities teacher can indicate in advance will be based on his total, and (let us hope) still-growing, view of the virtue of the subject he is teaching. Thus, objectives in the study of literature are the hypothetical sum of the possible desirable functions of literature; objectives in the study of history are the sum of the possible desirable functions of history. But these functions are multiple, and for any one person or group of people, not all of them can be equally operative at a given time. The teacher, however, is happy when he can infer that *any* of them are operative, and part of his expertise consists in being alert to the particular functions which appear, by the pupils' responses, to be viable at a particular time. Thus the teacher's opening moves in a lesson or course of lessons are necessarily experimental, offering a complex, *multi-faceted* object of study, and proposing a few selected ways of approaching it. He will be equally happy if his proposals are directly followed or if they lead, by reaction, to alternative but potentially fruitful ways of approaching the object of study; or even to the selection of a different object of study, so long as it is more or less within the field of his competence and confidence. He will maintain a general framework of order in his goals by reassessing periodically the direction of development of a given group of pupils; and at these points he may benefit by using formal test techniques, or he may rely on inference from recently marked written work or from comparison of impressions with other teachers.

Overspecificity and vagueness

If it is granted that the view advanced here has some substance, the ways in which the curriculum evaluation approach is inappropriate to the humanities can be fairly precisely identified, and we may then be able to proceed to develop a mode of evaluation more congenial to actual practice in these fields. As mentioned, curriculum evaluation takes its bearings from mathematics or science and then extrapolates to all subjects. Mathematical and scientific examples are therefore much more common in the literature, but if we focus on those humanities examples that are to be found, we observe the two apparently opposite weaknesses of over-specificity and vagueness. Thus Bloom's[10] examples in arts subjects are highly specific but paralyzing in their rigidity. In order to isolate different kinds of reasoning power it proves necessary most of the time to exclude opportunity for the pupil's inventiveness, expression of his own opinions or reference to his own experience. Can valid testing, let alone teaching, take place when motivations fundamental to the humanities are inhibited in this way?

Bloom does have sections designed specifically to test whether a student can 'achieve a novel organization of ideas'. For example, at one point[11] the student is required to produce an essay on the subject of private property. The instructions specify all the aspects of the subject which must be covered, set out the moral questions to be dealt with, demand to know 'the social effects of the kind of ownership which you favour or wish to attack' and the 'application of your theoretical position to one or more examples of property rights'. The directions for the essay themselves amount to about 600 words. It is hard to believe many humanities teachers would accept that a novel organization of ideas could be achieved in this way. Is there not a logical fallacy in the very fact of *prescribing* in such detail for *novelty?*

Basically what Bloom does is to split up all the complementary and multiply-interrelated mental functions *any* of which the alert teacher will be seeking to maximize at virtually *any* time, for any or all of them are likely to be involved in any reasonably complex learning activity such as reconstructing a historical event, appreciating a work of literature or discussing a social issue.

On the other hand, recent researches involving the curriculum objectives approach in humanities have, logically enough, used teachers' own stated objectives as their starting point, but have consequently been dogged by vagueness. Taylor[12] reports how the stated objectives of a group of English teachers, among others, were subjected to various forms of scrutiny, leading eventually to the judgment that they were very vague. Gronlund[13] prescribes these objectives for the interpretation of literature—

> Identifies the major and minor themes; states the author's purpose or message; identifies the tone or mood; explains why the characters behave as they do; points out specific parts of the literature that support the above interpretations; relates the literary work to other writings.

—which, though seemingly fairly specific, could apply to any level of interpretation of any work of literature. Eggleston and Kerr[14] obtained the following objectives from history teachers—

> (For the pupil to attain) an understanding of events, particularly of the relationship between cause and effect, with some appreciation of motivation. . . . Seeing the relevance of history both to present day events and to the locality in which the pupil lives. . . . Understanding the views held by people in other places, and an awareness of the world into which the pupil is entering.

—and then report that 'the unspecific nature of the objectives as they were originally phrased made it difficult to translate them into concrete test situations'. This did not lead, however, to a reappraisal of the function of objectives in the teaching process but to a substitution of 'something like the following process: firstly, devise a test item . . . which, it is agreed, tests something worthwhile; secondly, inspect the finished item and agree on what ability it tests; thirdly adopt the agreed ability as a teaching objective'[15]. Thus here test technique comes to dominate goals completely—a very unhealthy situation, one might think.

It is clear that what we have in all three cases are not behavioural objectives but general *aims*, i.e. descriptions of the approximate scope of the subject. What this approach shares with Bloom, despite this major difference, is the simple conception of teaching purpose referred to in the first sentence of this paper: the idea that intention wholly precedes and governs action, or should do. Ask teachers what their purpose is and they will give a statement of general aims. It does not necessarily follow that teaching will be improved by turning these aims into specific behavioural objectives or, where this appears impossible, by substituting such objectives independently arrived at. Nor does it follow that this procedure is the only means of evaluating the success of the teaching that then takes place.

Observation of teachers' 'assessment acts'

It is clear that in the course of the actual classroom encounter (and in marking written work) the teacher has a multiplicity of ways of expressing his general aims in individual assessment acts. That is to say, his aims serve to generate a potentially infinite number of specific evaluations of learning success on the part of the pupils, but can only do so in the process of recognizing beneficial learning outcomes *as they emerge*.

It follows that if we wish to analyse the real function of teaching objectives we must observe and analyse the actual multiplicity of assessment acts performed by the teacher in the course of his work. Mrs. Cameron-Jones has made an illuminating beginning in this region by producing a detailed categorization of assessment acts in the classroom[16]; but the even harder task of relating such acts to the subject matter of lessons remains to be attempted. The 'classroom interaction' branch of research[17], which involves the systematic observation of teacher-pupil interactions, has produced a number of schemata which might be important in such an endeavour. This field of research has so far been primarily concerned with the *frequency of occurence* of various teaching 'behaviours' (e.g. praise, criticism, acceptance of pupils' ideas). Also, the difficulties of systematizing its observations have tended to force it to focus mainly on the more traditional, static, teacher-directed type of lesson. While a relationship between frequencies of certain teaching behaviours and certain pupil 'outcomes' has been established, forms of systematic observation which could relate teaching 'behaviours' to lesson content and objectives (whether prescriptive or emergent) are still a long way off. It may be, indeed, that a fully systematic method for detailed observation of such a complex activity as teaching is a chimera, and that enlightenment is more likely to arise from other research strategies mentioned in passing above.[18]

We need definitions of objectives in the humanities not in order to prescribe teaching methods as if we were starting from scratch but in order to optimize the beneficial functions of the teaching which already takes place. We need them also to permit us to evaluate success in teaching these subjects *publicly*—to bring to light the types of evaluation which good teachers already employ as an integral part of their teaching, and hence to be able to demonstrate what this sphere of education achieves; we need them, finally, in order to guide the redesign of teacher education courses for optimum effectiveness.

The pertinent questions

Hopefully we are now in a position to identify some major questions about objectives in the humanities which actually derive from the nature of these subjects rather than from an overgeneralized analogy with the science subjects in which curriculum evaluation has found its primary models. Hitherto, curriculum theorists appear to have thought it legitimate to bypass the question of the function of a given subject, assuming this to have been decided satisfactorily sometime during the teacher's training, and to be embodied securely within his expertise. But it is likely, and indeed desirable, that this fundamental question has *not* been finally decided by the teacher: that his view of the function of his subject is tentative and permanently open to modification—especially by what more he discovers about his subject in the process of teaching it.

The outstanding 'structural' question about objectives in the humanities has already been implied in the foregoing critique: how do teachers in practice, distinguish appropriate *levels of complexity* for treating multi-faceted, multi-level subject matter? What do they expect of a good sixth-form essay on the character of Macbeth or the causes of the French Revolution, or from a drama improvisation or a discussion on freedom that they would not expect from a third form essay on the same subject? Apart from the observation of teachers' assessment acts already indicated, ways of approaching this question might include: (a) 'philosophic-sympathetic' dialogue with teachers, beginning by asking them their aims, as in the studies by Taylor and Eggleston and Kerr mentioned earlier, but transcending the inevitably vague initial statements by further questioning to draw out the more specific criteria likely to be in the backs of the teachers' minds; (b) observation of the different treatments given to similar subject matters with different classes; (c) analysis of teachers' assessments of written work.

Further questions crucial to objectives in the humanities might be divided (theoretically—*not organizationally*) into those concerned with basic knowledge and skills and those that are bound up with cultural and social values. The first kind are less problematic and might to some extent be amenable to classic curriculum evaluation methods. This area would include: (a) *information*: historical, social, geographical, civil rights, current affairs, conceived as the minimum basic knowledge necessary to picturing our world; (b) *articulacy and communication*: practice in sustaining debate, marshalling evidence, using a wide vocabulary, mastering different modes of writing and speaking appropriate to different purposes and milieux[19]; (c) *comprehension* of prose and discourse; (d) affective, in the limited sense of inducing an emotional disposition that makes intellectual work possible: a confidence in one's own judgment, receptivity to ideas and impressions, reciprocity in argument and relationship.

The central and most problematic aspects of the humanities are those that are bound up with social and cultural problems in general: creative expression, aesthetic appreciation, emotional development and moral and social values. And these aspects would be present even in subject matter introduced primarily with the more basic objectives in mind. We must beware of the illusion, noted above in Bloom, that exercises and tests could be designed for the 'low-key' objectives which did not also depend on the activation of the 'high-key' objectives.

Clearly we cannot make it a condition of formulating objectives in the humanities that we shall first have solved the cultural problems of contemporary society. Yet we must be aware of contemporary cultural development and debate as it forms the context, for better or worse, in which the cultural activity of the school takes place. Apart from the problem, already indicated, of defining progressive levels of complexity in the understanding of literature, history, art, religion and society, three major dilemmas stand out in this area: (a) how to evaluate creative writing and activities; (b) how to discover the educational effect of encounter with the 'high culture' in school; (c) how to discover the educational effect of utilizing elements of contemporary culture, especially 'popular' culture, in school.

Creative activities

The problem of how to evaluate creative writing and activity arises from the fact that while such practices are now fairly widely established in schools, there has been very little investigation of their exact function. It is possible that external evaluation of creative writing is relatively spurious; that it is an activity satisfying to pupils in itself, and with inbuilt 'organic' criteria of progress—i.e. that the benefit of improved internal organization of concepts and images is experienced by the pupil himself as an immediate improvement in conceptual fluency or in the apprehension of reality. This possibility would be well worth investigating. But a number of surrounding problems remain. Creative writing, like all other school tasks, takes place within a framework of teacher assessment. The teacher marks the work or assesses it more informally; the teacher also has the function of proposing the next piece of work, which is presumably in part dependent on the level of attainment visible in the last; also, pieces of work may be submitted as course-work for CSE or other examinations. It appears then that creative work is subject to external assessment whatever its internal benefits may be. But the criteria of such evaluations are usually very inexplicit. The emphasis of this particular type of innovation has been largely on the *stimulating* of creative writing, discussion and drama.[20] It would be valuable to discover what criteria are operative in teachers' informal assessments of creative activities. It might be hypothesized that where imaginative writing or drama is concerned, the implicit criteria have some relation, however tenuous, to literary critical criteria in general; or, in other words, that there is a continuum between criteria for pupils' creative work and, on the other hand, the criteria invoked in relation to the high culture.

Culture, traditional and contemporary

Regarding 'high culture', that is the best of art, literature and music from the Western tradition, it is remarkable that its educational effect, other than on the ability to appreciate it *per se*, has been so little investigated. The benefit of contact with it has been a presumption resting on the *intrinsic* value attributed to it by critics over the generations. But just how this intrinsic value interacts with the perceptions of adolescent minds is still something of a mystery. Related questions are: does education actually facilitate an encounter with high culture or is it more truly based on an encounter with critical opinions *about* high culture?[21] If so, what is the educational effect of the encounter with received critical opinion? Further: can all material drawn from the high culture be presumed to be equally suitable for educational purposes? Are there attitudes and values embodied in the high culture which may be objectionable on contemporary moral or social grounds (glorification of war, subservience of women, inferiority of races)? Or does the definition 'high culture' itself imply acceptance only of works embodying a universal humanitarianism? If the latter, does this valuation correspond to works actually used in schools and the way they are presented? Finally, is there a tendency to use *minor* examples of the high culture with younger age groups or 'less able' pupils because these examples are regarded as more accessible; and if so, do the works so selected genuinely share the virtues attributed to the major works?

The problems associated with the utilization of contemporary culture in schools are similar to these but from the opposite angle. In contrast to the generally positive assumptions about the educational value of the high culture, misgivings are frequently expressed about the effects of popular culture in general and the wisdom of using elements of it in school. Its unavoidability in daily life is generally accepted, but many educationists[22] see the function of school as being specifically to counteract its influence. Like 'high culture' the term 'popular culture' or 'contemporary culture' clearly covers a wide variety of material, and it is also possible to treat this material in a variety of different ways within the school, whether purely as a stimulus, as a stepping stone towards high culture, as an object of criticism or as legitimate cultural material in itself. The questions again are: just how is it used in school and what are its educational effects?

There have been various studies of pupils' degrees of appreciation of literature, and the Schools Council has promoted investigations of reading habits and preferences[23] and cultural dispositions[24] among pupils. But generally, research into literature and art has meant research into the influences *on* literature and art, not the influence of literature and art on people, just as research into history has meant research into historical processes, not the influence of learning about those processes on the world view or social behaviour of pupils. In short, the educational value of the basic materials of the humanities is usually regarded as implied in their 'intrinsic' value, so that researches into the teaching and learning of these subjects have almost invariably been investigations of how far, how well, how quickly pupils have been able to produce approved critical comment or appreciation.

Discursive writing on the subject has often invoked wider criteria, claiming that the study of literature enhances relationships, that the appreciation of art aids emotional fulfilment, that participation in creative drama facilitates social adjustment, that the study of the humanities in general stimulates the individual's ability to coordinate his subjective impressions into intelligible and communicable structures. Undoubtedly these wider criteria are espoused, often with enthusiasm, by teachers; but their effect is not generally appreciated in any concrete way outside the subjects themselves. It is perhaps particularly important at this time that humanities teachers should develop their objectives and criteria of assessment in a reasonably public manner in order to forestall any danger of the imposition of objectives and criteria arbitrarily extrapolated from scientific subjects, as campaigns for demonstrable educational efficiency gather momentum. At the same time, since humanities questions are the value-questions of our society in general, the development of more explicit objectives in the humanities might throw some light on the fundamental questions of social aim by which our society is bedevilled and to which efficiency alone is no answer.

Notes

1. For example S. WISEMAN and D. PIDGEON, *Curriculum Evaluation*, Slough: NFER, 1970; and J. F. KERR, (Ed.), *Changing the Curriculum*, London: University of London Press, 1968.
2. Evidence of the almost continuous 'state of emergency' in which primary school teachers work will be found in S. HILSUM and B. CANE, *The Teacher's Day*, Slough: NFER, 1971 (short version: *The Teacher at Work* by S. Hilsum, Slough: NFER, 1972). The NFER is currently conducting a complementary inquiry into the secondary teacher's work.
3. From DOUGLAS A. PIDGEON, 'Individual freedom and the reform of teaching' in M. ASH, (Ed.), *Who Are the Progressives Now?*, London: Routledge and Kegan Paul, 1969.
4. See for example C. P. ORMELL, 'Educational objectives—Bloom's taxonomy and the problem of classification', to be published in the next issue of this journal, *Educational Research*, 17, 1, November, 1974.
5. See M. PARLETT and D. HAMILTON, *Evaluation as Illumination: A New Approach to the Study of Innovatory Programs*, Edinburgh: Centre for Research in the Educational Sciences, Occasional Paper 9, 1972.
6. See for example DES, *Educational Priority, Vol 1, EPA Problems and Policies*, by A. H. HALSEY, London: HMSO, 1972.
7. A notable recent example of participant observer research, though not in the classroom, is James Patrick's *A Glasgow Gang Observed*, London: Eyre Methuen, 1973. Patrick spent periods of time seeing gang life from the inside, and discusses problems of objectivity and authenticity in this approach, in the course of reporting his findings.
8. See for example MARILYN M. NICKSON, *Educational Technology*, London: Ward Lock Educational, 1971.
9. See GABRIEL CHANAN and LINDA GILCHRIST, *What School is For*, London: Methuen, 1974, especially Chapter 6. (In press).
10. B. S. BLOOM, *et al.*, *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*, New York: Longmans, 1956.
11. Ibid. p. 177.
12. P. H. TAYLOR, *How Teachers Plan Their Courses*, Slough: NFER, 1970.
13. N. E. GRONLUND, *Stating Behavioural Objectives for Classroom Instruction*, London: Macmillan, 1970.
14. J. F. EGGLESTON and J. F. KERR, *Studies in Assessment*, London: English Universities Press, 1969, p. 78.
15. Ibid. p. 82.
16. M. CAMERON-JONES and A. MORRISON, 'Teachers' assessments of their pupils', in G. CHANAN, (Ed.), *Towards a Science of Teaching*, Slough: NFER, 1973.
17. A selection of short accounts and examples of classroom interaction research will be found in

A. Morrison and D. Macintyre, (Eds.), *Social Psychology of Teaching*, Harmondsworth: Penguin, 1972.

18. See notes 5, 6 and 7.
19. See Harold Rosen, 'Written language and the sense of audience', *Educational Research*, **15**, 3, 1973.
20. See for example the Schools Council/Nuffield Foundation *Humanities Project*, London: Heinemann Educational, 1970.
21. C.f. Douglas Holly, 'Crypto-elitists and pseudo-radicals', *Times Educational Supplement*, 26th January 1973.
22. See for example G. H. Bantock, *Culture, Industrialization and Education*, London: Routledge and Kegan Paul, 1968.
23. Reported by Pat D'Arcy, *Reading for Meaning, 2: The Reader's Response*. London: Hutchinson Educational, 1973.
24. Reported by Graham Murdock and Guy Phelps, *Mass Media and the Secondary School*, London: Macmillan, 1974.

British Journal of Educational Psychology

Editor: J. D. NISBET, Department of Education, University of Aberdeen, Scotland

Vol. 44 Part 2 June 1974

The *British Journal of Educational Psychology* is concerned mainly with reporting the results of experimentally-based studies involving children and students and their education, together with more general questions about the psychology of learning and studies of social problems where there is a psychological emphasis to the enquiry. The Journal has a circulation of about 4,000 and has subscribers in eighty countries throughout the world. It is one of the oldest and best known journals in educational psychology in Europe.

It appears three times a year with about 120 pages in each issue covering half a dozen main papers (up to 4,000 words) and 6 to 10 research notes (about 2,000 words). There is also a Book Review section in which about 50 new publications each year are reviewed by the country's leading educational psychologists.

H. Goldstein and K. Fogelman: Age Standardisation and Seasonal Effects in Mental Testing.
C. McGeorge: Situational Variation in Level of Moral Judgment.
P. L. Gardner: Research on Teacher Effects: Critique of a Traditional Paradigm.
P. Taylor: Lecturers' Perceptions of the Influence of Ideas on the Curriculum of Colleges of Education.
A. R. Hakstian and R. B. Cattell: The Checking of Primary Ability Structure on a Broader Basis of Performance.
A. J. Malpas and M. Brown: Cognitive Demand and Difficulty of GCE O Level Mathematics Pre-Test Items.
J. G. Houston and A. E. G. Pilliner: The Effect of Verbal Teaching Style on the Attainment of Educational Objectives in Physics.
B. Wilcox: The Teaching of Serial Tasks Using Chaining Strategies.

Critical Notice:
G. E. R. Burroughs: Methods of Educational Enquiry (Open University).

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Establishing Children's Learning as the Criterion for Teacher Effectiveness

K. B. Start, Faculty of Education, University of Melbourne

Summary

The purpose of teachers, teaching, and the curriculum in schools is to promote learning. Attempts to clarify teacher effectiveness in relation to pupil learning where they exist in past literature have been notably ambiguous. Rather global measures by headteachers, college tutors and inspectors have been used instead, or dimensions have been extracted from detailed observations of classroom interaction. Reasons for recent concern with evaluation of the applicability of the teacher's training, method or lesson content to his task are advanced, and it is argued that once the dimensions of children's learning that are to be the specific responsibility of the school are defined, criteria for successful learning could be evolved by teachers along each dimension. Assessment and interpretations would, however, need to involve wide considerations of the concepts of learning, behaviour, interest, motivation etc., and this, although an enormous task, must be done.

There should be but one overriding purpose for teachers, teaching, schools and curricula—to evoke learning in children.

Children's learning should not be narrowly defined as cognitive nor limited to affective, social or vocational skills. Neither should it be lost in terms so broad as to be meaningless, such as personal growth; nor follow any particular idiosyncracy of teacher, philosopher or politician. Indeed it is about time that society thrashed out just what this learning should encompass. However, that is an issue of decision that can wait until another time. In this paper, children's learning is not seen as a narrow concept, but one the multiplicity of which remains capable of more detailed definition when a practical application is needed.

The research on teachers, teaching and teacher education has given rise to no notable theory of teaching. Neither has it been notable for the consistency of its findings other than that there seems to be little agreement between the many different criteria used to judge successful teaching (Wiseman and Start, 1965; Koskenniemi, 1965; Start, 1967). It has been bedevilled by the lack of an agreed definition of effective teaching, and of teacher competence (Getzels and Jackson, 1963; Tarpey, 1965; Lomax, 1972). It is truly amazing that so much research has occurred in an area where the essential criterion evades definition (Domas and Tiedman, 1950).

The link between quality of teaching and quality of learning has frequently been avoided. Observe, for example, the California Teachers Association Statement of 1956, which claimed that merit ratings were wrong as teacher effectiveness could not be evaluated, and that 'ratings of teaching ability would cause jealousies among teachers' (p.4). This position has been modified considerably in the past 20 years, with the growing work at centres such as Stanford on teachers and teaching methods in relation to the cognitive, social and personal development of children. Even so, the Stull Act (passed by the Californian legislation on 20.7.71) which makes it mandatory to find a measure of teacher efficiency in terms of pupil learning, has caused some turmoil.

In England the early payment-by-results system, its accompanying inspections, and concentration on the four Rs, was probably the first practical use of aversion therapy. Limited understanding of the impact on learning of economic, social and domestic issues made such assessments crude and often unjust. To survive, the teacher taught the limited curricula to distraction, and consciously or unconsciously the teaching profession resolved to remove the constraint and never to be put in that position again. But as seems to happen all too often, *today's innovation* to ward off yesterday's *rigidity* becomes tomorrow's *immutable doctrine*. In rejecting yesterday's limited aims, methods and measures, it seems that we have also hedged out the principle that children's learning is the sole purpose of compulsory education and hence must be the major criterion for assessing its effectiveness (unless we accept roles of child minding and delaying entry onto the labour market).

Of the many studies of teachers, teaching methods and teacher education, few have involved measures of children's learning. If we limit any review to studies in England the number becomes miniscule. What are used might be termed adjunct criteria. These appear to be logical derivations from what are seen to be 'obviously good' teachers and teaching practices. In this category are professional global assessments by experienced practitioners such as headteachers, college tutors and inspectors. They know that good teaching produces good learning, and 'know' they can recognize a good teaching performance when they see it. The current disenchantment with the grading of teaching practice might well reflect the lack of agreement between college tutors, as well as among teachers and headteachers, on what 'good' teaching is.

At another level are the classroom observation schedules with structured recording by which the activities of teacher and child can be categorized. These can be refined by the introduction of audio-visual or closed-circuit television methods, with or without self confrontation. The dimensions extracted from such observational schedules become ends in themselves and can then be used in teacher training to attempt to direct the behaviour of the young teacher into 'good' dimensions and modify his behaviour to make him 'better' at such activities. A further step away from the actual process of teaching is the use of attitude or personality scales, usually by group techniques, and the attempt to associate certain outcomes with philosophically 'good' teaching attitudes, e.g. child centredness, or with value dimensions on which differences can be shown to exist between different groups of teachers.

All these second-order criteria probably do relate to children's learning, but only in the way that by weighing someone you can estimate his height, because 'big' people are often tall.

One must ask why relatively distant criteria have been used instead of children's learning; for as Howsam (1960) said in California 'there is general agreement that the profession is not presently able to make use of the pupil gain criteria in measuring teacher effectiveness' (p.16). The backlash against the payment-by-results era certainly may be regarded as the major culprit. The success of British education ironically must also take its share of the 'blame'. In fact the traditional professional responsibility of the classroom teacher has been a hallmark of the British educational scene, and his judgement has been paramount. It is only with the recent explosion of the tasks that society has wished upon the teacher that the question of the applicability of this training, method or content to that task has been seriously raised. And in raising it the teaching profession has found itself ill prepared for hard evaluation. The pressure under which the educational system has enlarged must take some of the responsibility, as must the flood of innovation that has gone through the school systems. Too much too quickly hardly leaves time for quality control. In a period of careering change, innovation for innovation's sake is destructive (Toffler, 1973) and the need for evaluation of innovation and change is critical to ensure that only the efficient and the 'better' is acceptable. Even then, any improvement should be set against the cost in economic and in personal terms.

Two other reasons remain, and these could be as important as the aversion to payment by results. One is the technical difficulty of estimating the improvement resulting from the effect of the teachers, and the other is the psychological nakedness that would be felt by teachers faced with such a direct measure of their performance.

The technical difficulties partly consist of knowing just what is it that teachers are expected to do in school. Just what is it that parents rely on their children learning at school? Perhaps teachers have skills that should be focussed only on particular outcomes, not deployed over the whole range of human development. Certainly there are things that teachers can do best. Others they may well do better than other agencies, i.e. parents. If for no other reason than the length of the teaching day, a compromise should be achieved between the philosophy of school as a dumping ground for social tasks and as a limited, academic, abstract-thinking ivory tower. What society expects its children to learn at, from or through school must be discussed with the teachers and then defined in meaningful ways.

Once the dimensions of children's learning that are to be the specific responsibility of the school are defined and agreed by the teaching profession, the criteria for successful learning could be evolved by the teachers along each dimension, whether it be cognitive, social, religious or personal. Measures of progress on these dimensions could be evolved with greater or lesser accuracy depending on the clarity of the definition of the criteria. Before-and-after testing would seem the obvious, perhaps too obvious, method of estimating pupil gain, though the maturation effect of the teaching might require long-term assessment as the 'after' measure. But the biggest difficulty would be the interpretation.

Children's learning is not a simple, direct consequence of a teacher teaching them. Though it may not be socially or politically palatable, some children seem to learn more easily than others, whether through environment, heredity or both. Children's socio-cultural background can give them different material and attitudinal advantages and disadvantages in learning. Children's interests and motivations not only vary in quantity but in quality, and the interaction between the area of learning and the child's interest and motivation to learn can determine his achievement. Again our concept of learning must not be limited to the one style of abstract cognitive thinking, but embrace all dimensions of human behaviour, social interaction and skills. The greatest tasks of definition, measurement and evaluation lie in these directions. The sincere belief of experienced teachers convinces us that whereas a 'warm' classroom atmosphere enhances learning, the teacher's hair colour, weight or height has no effect. But whether autocratic teaching is better than democratic teaching for independent and dependent children's learning, both cognitive and affective, is a question that would be better resolved by direct assessment than indirect opinion.

Education is more than the learning of facts. Often in the past, in schools of narrow outlook, this limited cognitive development may have become overdeveloped, as did the premium put upon academic-content examinations. Nowadays the realization of the breadth of the concept of what education is all about is in full flood, and narrow academic limits are rejected along with the measures with which they were associated. Is it not time we checked that the over-emphasis and over-precision in one (cognitive) dimension has not been replaced by an equally negative generalization and quantification-phobia? By all means let us multiply the aims of the schools, curricula, teaching and teachers, but surely one must want to know whether these aims are being achieved.

Recently a study by the author asked nearly 1,000 teachers to list the five ways they would be able to tell a good teacher from a weak teacher. Of the 5,000 possible responses, 37 mentioned whether children learned. Hence less than 4 per cent of teachers raised the direct criterion. Perhaps it was implicit in the other teachers' thinking, but it suggests just how far we have moved from overt consideration of the teacher's purpose in school—to effect children's learning.

It will be a formidable task to devise ways of assessing pupil learning along many dimensions. It is a daunting prospect to attempt to tease out the contributions that teaching had made in each of these dimensions. To allow for the quantity and quality of prior knowledge, experience, learning strategies and attitudes is a task-and-a-half by anyone's standards. But it must be done. We must know what efficient teaching and teachers are. Without such knowledge how can we pretend to train student teachers? Presumably one aims to produce efficient rather than inefficient teachers. Certainly one should question any training that has random outcomes. If we could establish what are good teaching and teachers, and produce courses to develop these methods, attitudes and people, selection for teacher training would become meaningful. Successful candidates would have, in a developed or latent form, those attributes that would maximize the benefit of the training.

It is a long, involved and meticulous task that has to be done, but it cannot be shrugged off because of that. Equally complex tasks have been attempted. It is a matter of history that studies over decades have been undertaken and sums of a quarter of a million pounds have been granted for similar purposes. Even within teacher education, certain instruments such as Flanders' Interaction Analyses have evolved over 10 years or more. So there are precedents for large, complex studies of formidable issues. Why then do we avoid the crucial criterion in teacher effectiveness—pupil learning?

References

CALIFORNIA TEACHERS' ASSOCIATION (1956). 'Information on the merit rating of teachers', *Research Bulletin*, 98, December. San Francisco: C.T.A.

DOMAS, S. J. and TIEDMAN, D. V. (1950). 'Teacher competence: an annotated bibliography', *J. Exp. Educ.*, 19, 101–218.

GETZELS, J. W. and JACKSON, P. W. (1963). 'The teacher's personality and characteristics', in GAGE, N. L., *Handbook of Research on Teaching*. Chicago: Rand McNally.

HOWSAM, R. B. (1960). *New designs for research into teaching competence*. Project report for Joint Committee on Personnel Procedures (California School Boards' Association and California Teachers' Association). Burlingame, California.

KOSKENNIEMI, M. (1965). *The Development of Young Elementary School Teachers*. Helsinki: Academie Scientiarum Fennicae.

LOMAX, D. E. (1972). 'A review of British research in teacher education', *Review of Educational Research*, 42, 3, 289–326.

START, K. B. (1967). *A follow-up of the 1961 group of teachers*. Mimeographed Report No. 1. School of Education, University of Manchester, to Department of Education and Science.

TARPEY, M. S. (1965). 'Personality factors in teacher training selection', *Brit. J. Educ. Psychol.*, 35, 2, 140–9.

TOFFLER, A. (1973). *Future Shock*. London: Pan Books Ltd.

WISEMAN, S. and START, K. B. (1965). 'A follow-up of teachers five years after completing their training', *Brit. J. Educ. Psychol.*, 35, 342–61.

NFER

COMPUTER SERVICE

The National Foundation for Educational Research has now installed its own computer—an IBM 1130.

We anticipate a certain amount of spare time in the next year or so, and are prepared to make the computer available to Educational Research Workers at favourable rates.

Card-punching will also be undertaken, at competitive rates.

Details

A comprehensive range of programs is available. These include regression analysis, factor analysis (up to 100 variables), a general tabulation program, χ^2 analysis, and analysis of variance. In addition, there is a wide range of data processing programs for those researchers who need to handle a large amount of data.

Rates

Computer: £21 per hour.

Data Processing: £27 per 1,000 cards.
Punched and Verified.

Specimen Costs:
Regression analysis of 8 variables on 40 subjects—£5: 15 tables from 2,600 subjects—£10.

All charges subject to 10% VAT.

Telephone Slough 28161, or write to
A. N. James,
NFER, The Mere, Upton Park, Slough, Bucks.

Conceptions of School Among Pupils Affected by the Raising of the School Leaving Age

Alan Smithers, George Avis and David Lobley,
School of Research in Education, University of Bradford

Summary

From among the entire third-year (13-14 age range) of four comprehensive schools in the West Riding those pupils who wished they could leave at 15 and those who wished to stay on till they were 18 were identified. There were similar proportions of boys and girls in the two groups, but the 'leavers' were predominantly from manual backgrounds and were generally judged by their teachers to be unlikely to pass public examinations. The two groups differed in their conceptions of school, and there were differences with sex and ability, but there was also a considerable measure of agreement: the school tended to be seen in mainly instrumental terms such as helping pupils do well in GCE and CSE examinations and qualify for good jobs; expressive functions such as cultural development and education for leisure tended to be de-emphasized. It was the would-be leavers and those who were not rated academically able by their teachers who tended to attach more importance to the expressive role of the school. In the prevailing examination-dominated climate this may have been a rationalization of their lack of academic success. If the extra year at school is to be an educational advance, the problem of appearing to cater for second-best must somehow be solved.

Introduction

The raising of the school leaving age has come at a time when increasing numbers of school children are choosing to stay on at school voluntarily. Between 1966 and 1971 the proportion who chose to stay on in maintained schools rose from 41·6 per cent to 57·2 per cent (DES, 1972). Yet ROSLA also highlights the persistence of the phenomenon of early leaving which received much attention in major national inquiries of the recent past (e.g. CACE, 1954; CACE, 1959; Committee on Higher Education, 1963). Such studies emphasised the strong association between early leaving and social-class background, and underlined the consequent wastage of talent. In a more recent large-scale investigation Morton-Williams and Finch (1968) showed that some 80 per cent of the statutory-age leavers among their random sample of pupils in maintained schools in England and Wales were the children of manual workers.

An early study by Fitt (1956) indicated that children from a working class background generally held a less favourable attitude to school than those from a middle class background. Attitude to school has also been shown to be related to ability and sex. A positive correlation between ability and attitude to school was reported by

---

Acknowledgments

The authors are grateful for the help of R. Busby, R. J. Casken, A. J. Coward, Mavis Curtis, A. Glendenning, M. J. O. Mahoney, R. B. North, G. Taylor and Sheila Dann. This work was carried out under the direction of the senior author as part of the Advanced Diploma in the Psychology and Sociology of Education, University of Bradford, 1971-73.

Hargreaves (1967) who noted a significant decline in satisfaction with school life from upper to lower streams in the fourth form of a boys' secondary modern school. Similarly, Halsall (1971) found more favourable attitudes to school in the top stream of the fourth form in a comprehensive school, but the attitude of all pupils regardless of stream was somewhat unfavourable. Fitt (1956) found that girls tended to like school more than boys, but, although this has subsequently received support at the primary stage (Wisenthal, 1965; Sharples, 1969; Barker Lunn, 1969), the evidence for the secondary level is conflicting (Allen, 1960; Carter, 1962; Kniveton, 1969; Halsall, 1971).

Following the Government's decision in 1964 to raise the school leaving age, several studies have investigated the attitude of pupils to the aims and content of secondary school education. The most important of these studies was an extensive survey conducted by the Government Social Survey for the Schools Council (Morton-Williams and Finch, 1968). But two smaller studies which preceded it are of some interest. Among a sample of over a thousand 14-year-olds in Welsh secondary schools, Lotwick (1965) found a keen awareness of the world of work and 'a strong desire on the part of most pupils to improve their qualifications and their readiness to tackle, in an effective manner, the job of their choice'. Another survey carried out in the Midlands by Clark (1967) examined the views of 447 pupils of average and below average ability in the fourth year of 13 non-selective secondary schools as to what they thought they should be doing in the extra year of schooling, and how they felt they should be treated. A most interesting feature of this study was the wide difference in response between the sexes. A small majority of the pupils were opposed to the idea of raising the school leaving age, but girls significantly more than boys. Despite an overall proportion of 62 per cent who said that they would not remain on at school unless they were compelled to, Clark (1967) felt that his study indicated 'a bigger and more positive fund of goodwill toward school, and particularly toward external examinations than many teachers would give credit for to children of this ability range'. The three most popular subjects or activities for the extra year were: outdoor pursuits; visits to factories, offices, etc.; and parentcraft, sex and health education. Then came practical subjects, science and mathematics; apart from English, little weight was given to the traditional humanities. Girls were significantly less enthusiastic about science and mathematics, but placed parentcraft, sex and health education, as well as speech training, significantly higher than boys.

Broad similarities exist between these findings and those of the Schools Council national survey (Morton-Williams and Finch, 1968). A sample of 4,618 pupils and ex-pupils aged 13-16 years were asked in an interview to rate 27 possible school objectives. In general, the authors found 'very wide support given to what has been called the instrumental role of the school, that is to those aspects which provide the keys to success in later life—giving pupils the knowledge and skills which would enable them to obtain as good a job or career as possible; helping them to deal with the practical aspects of everyday living, managing money, running their homes . . .', but on the other hand 'the expressive functions of the school, providing interests and enjoyment, recreational or cultural, giving outlets for energies and emotions, broadening the mind and increasing awareness of the world, were rarely seen as of importance'.

It should be remembered that these studies were carried out over six years ago, in 1966, and that they examined the views of pupils who were not *actually* to be affected by ROSLA. This paper attempts to assess what sort of attitudes were held in 1972 by pupils who faced the reality of a compulsory further year of schooling.

Methods

The sample consisted of the third year (13-14 age range) of four comprehensive schools in an industrial area of the West Riding. These pupils (300 boys and 481 girls) are among the first cohort who will be obliged to stay on till 16, following the raising of the school leaving age in 1972-3. Three of the schools were co-educational, and one an all-girls school. One of the mixed schools was a 'middle school' (11-14); the rest were 'all-through'. None of the schools uses rigid streaming, but there is 'banding' into three groups: those being prepared for GCE examinations, those being prepared for CSE examinations, and those considered to have no chance of passing public examinations. The curricula of these three groups differ somewhat, with, for example, the third group taking remedial English and maths instead of a second language and more science.

The children were asked to rate 27 possible

Table 1: 'Leavers' and 'stayers' by social class and ability band (significant differences only)

ABILITY BAND	PER CENT *Social Class**		
	I + II	III	IV + V
'Leavers'			
GCE	58·4	18·4	12·8
CSE	25·0	46·0	20·6
Not of examinable standard	16·6	35·6	66·6
N	12	76	39
'Stayers'			
GCE	79·3	65·6	48·7
CSE	19·5	27·8	32·4
Not of examinable standard	1·2	6·6	18·9
N	82	137	37

* Registrar-General's categories; insufficient information was available to place some children.

school objectives (see Appendix) taken from the interview schedule of Morton-Williams and Finch (1968) on a three point scale—'very important', 'fairly important' or 'not important'. Additional questions were included to ascertain sex, father's occupation, and preferred age of leaving.

Of the total sample of 781, 175 said 'yes' in reply to the question 'If you could still leave school at 15, would you?' These have been designated the 'would-be leavers', or 'leavers' for short. Of the 506 who said 'no', 287 indicated '18' as the age that they thought they would be leaving school, and they have been designated the 'stayers' (the other 319 mostly expected to leave at 16).

Comparisons have been made between the 'leavers' and 'stayers', and responses have also been analysed by sex, social class and ability band. There were very few differences between the schools: with sex controlled for, only in four of the 243 comparisons were differences significant beyond the one per cent level which is not very different from what you would expect by chance. The data have therefore been aggregated.

Results

Equal proportions of boys and girls were found among the 'leavers' and 'stayers', but there were highly significant relationships between desire to stay on, and social class and ability band. As Table 1 shows the lower the social class and the lower the presumed ability, the more likely the pupil is to want to leave early.

When the objectives of the 'leavers' and 'stayers' were compared, it was found that they differed significantly on seven of the 27 items. As Table 2 shows, the 'leavers' were more interested in learning about job conditions by visits to work places, and running a home than the 'stayers'. They also placed greater emphasis on the school organizing clubs out of school hours. In contrast, the 'stayers' were more interested in being helped to examination success and a good job, in being helped to become independent and learning about what is going on in the world nowadays.

These differences, however, occurred within a broad framework of similarities. For most children, whatever their leaving preferences, the real business of the school was helping you to examination success, teaching you things that will help you to get a good job and teaching you to manage your money. Less than a quarter of the pupils thought it very important the school should provide opportunities for enjoying drama and poetry, should organize outings and visits to places of interest, or should help develop leisure time interests.

The responses of the 'leavers' and 'stayers' were analysed by social class, ability band and sex. It was found that only one item significantly distinguished the social classes. Among the would-be leavers 54 per cent of those from the Registrar-

General's classes IV and V rated being taught to speak well 'very important' compared to only 29 per cent of those from class III (there were only 12 would-be leavers in classes I and II). No significant differences were found among the 'stayers', but only 37 were from classes IV and V.

Ability band appeared to be closely associated with the pupils' conception of school objectives. As Table 3 shows there were comparatively few differences between the GCE and CSE groups, but the 19 who wished to stay on to the sixth form but who were classed by their teachers as non-examinable were a distinctive group. They tended to emphasize the expressive functions of the school, e.g. taking pupils away on holidays, organizing visits to places of interest and outings to theatres, and teaching a range of subjects for general interest. They also rated developing personality and learning how to speak well as more important than the GCE and CSE groups. In contrast, the examinable groups appeared more instrumental in stressing the importance of learning things relevant to good jobs and managing money.

Interestingly, as Table 4 shows, a very similar pattern was found among the 'leavers' with the emphasis among those classed as non-examinable again being on expressive non-academic functions of the school; learning to speak and write well were also rated highly.

When the sexes were compared (Tables 5 and 6) it was found that girls were significantly more interested in learning to run a home, manage money and developing self-confidence. The boys placed more emphasis on hobbies. A surprisingly high proportion of the female leavers rated being helped to pass examinations 'very important' even though they wished they could leave before sitting any public examination.

Table 2: School objectives of 'leavers' and 'stayers' compared

Objective†	Per cent 'very important' Leavers (N=175)	Stayers (N=287)	χ^2
1. independent	64	75	5·55*
2. exam success	84	98	33·15***
3. personality and character	38	31	2·32
4. directly relevant to job	77	75	0·31
5. world affairs	18	35	14·52***
6. range of subjects	29	32	0·48
7. good job	86	93	5·55*
8. morality	43	39	0·43
9. speech	37	35	0·15
10. writing	39	48	2·30
11. drama	6	10	1·42
12. poetry	10	5	2·14
13. running a home	74	58	12·39***
14. managing money	77	76	0·01
15. sex education	51	46	0·98
13. self-confidence	45	51	1·76
17. appearance	40	32	2·74
18. social skills	45	49	0·34
19. hobbies	23	16	3·38
20. career knowledge	67	74	1·60
21. job conditions (hrs.)	42	41	0·03
22. job conditions (visits)	61	51	4·07*
23. outings (theatre etc.)	23	21	0·17
24. visits to places of interest	22	18	1·14
25. clubs	41	28	7·73**
26. holidays	39	36	0·40
27. residential courses	38	45	1·48

† see appendix for items in full.

* significant beyond 5 per cent level, ** significant beyond 1 per cent level, *** significant beyond 0·1 per cent level.

Table 3: School objectives of 'stayers' by ability

Objective	Per cent 'very important' GCE (N=198)	CSE (N=70)	NE (N=19)	χ^2 GCE *v* CSE	GCE *v* NE	CSE *v* NE
3. personality and character	27	36	53	n.s.	5·29	n.s.
6. range of subjects	29	32	69	n.s.	12·21	8·57
7. good job	95	92	74	n.s.	8·03	n.s.
8. morality	37	39	63	n.s.	6·05	4·58
9. speech	30	42	63	n.s.	8·66	n.s.
14. managing money	78	79	53	n.s.	5·94	5·11
17. appearance	25	53	32	18·52	n.s.	n.s.
22. job conditions (visits)	43	64	79	8·99	8·79	n.s.
23. outings (theatre etc.)	19	20	47	n.s.	8·11	5·66
24. visits to places of interest	13	23	47	n.s.	15·02	4·44
26. holidays	31	44	63	4·17	8·13	n.s.

Critical values of χ^2 for d.f. = 1, $P<0·05$, 3·84; $P<0·01$, 6·64; $P<0·001$, 10·83.

Table 4: School objectives of 'leavers' by ability

Objective	Per cent 'very important' GCE (N=29)	CSE (N=68)	NE (N=78)	χ^2 GCE *v* CSE	GCE *v* NE	CSE *v* NE
1. independent	83	66	56	n.s.	6·34	n.s.
6. range of subjects	17	12	47	n.s.	6·85	19·59
9. speech	21	22	55	n.s.	9·11	14·62
10. writing	31	28	51	n.s.	n.s.	7·32
17. appearance	21	48	40	5·45	n.s.	n.s.
24. visits to places of interest	14	15	32	n.s.	n.s.	5·08
25. clubs	14	44	49	7·19	9·40	n.s.
26. holidays	10	53	39	14·05	6·57	n.s.

Critical values of χ^2 for d.f. = 1, $P<0·05$, 3·84; $P<0·01$, 6·64; $P<0·001$, 10·83.

Table 5: School objectives of 'leavers' by sex

Objective	Per cent 'very important' Boys (N=74)	Girls (N=101)	χ^2
1. independent	54	72	5·43
2. exam success	66	93	18·85
3. personality and character	28	46	4·35
13. running a home	66	80	4·17
14. managing money	66	85	7·64
16. self-confidence	34	53	5·31
17. appearance	23	53	14·72
19. hobbies	32	16	6·02

Critical values of χ^2 for d.f. = 1, $P<0·05$, 3·84; $P<0·01$, 6·64; $P<0·001$, 10.83.

Table 6: School objectives of 'stayers' by sex

OBJECTIVE	PER CENT 'VERY IMPORTANT' Boys (N=106)	Girls (N=181)	χ^2
13. running a home	43	67	14·83
14. managing money	63	85	15·95
16. self-confidence	42	57	6·59
19. hobbies	24	11	7·71
27. residential courses	32	52	10·28

Critical values of χ^2 for d.f. = 1, $P<0·05$, 3·84; $P<0·01$, 6·64; $P<0·001$, 10·83.

Discussion

The present research again demonstrates the strong link between social-class background and early leaving, but, interestingly, when stated leaving preference was controlled for, no relationship between social class background and conceptions of school was detected. In part, this is probably due to the nature of the comparisons that were possible—only 9·4 per cent of the 'leavers' were from the Registrar-General's social classes I and II and only 14·4 per cent of the 'stayers' from social classes IV and V. Although no within-group differences were demonstrated, there were considerable differences between the groups. Thus the 137 'stayers' and 76 'leavers' who were from social class III had come to hold rather different views of the school's function.

Sixty-five per cent of the 'stayers' as opposed to only 18 per cent of the 'leavers' were adjudged capable of passing GCE examinations. It is not clear how accurate these assessments were, but, reinforced by differences in curricular provision, they seem to have been accepted as realistic by the pupils who had decided that school could do little more for them. Nevertheless these pupils continued to place considerable emphasis on examination success as a school objective.

The pupils adjudged not to be capable of passing examinations but who wished to stay on are an interesting group. Everyone of the 19 rated examinations success as a very important school objective. However, they also tended to give more weight to general interests and character development than the other pupils. They may be a very depressed group—valuing exam success very highly but not being thought to be capable of it.

In contrast, there were 29 pupils adjudged capable of passing GCE examinations, and 68 CSE examinations, who wished to leave without sitting them. Here the school itself may have been a source of frustration or there may have been outside pressures.

Comparison of the profiles of the boys and girls shows that the sexes continue to be socialized differently. The girls tended to rate running a home, managing money, developing self-confidence and personality, and looking after appearance more highly than the boys. It might be that, particularly for early leavers, the last two would be more important for girls in getting a good job but there still appears to be conformity to the traditional woman's role.

Although, so far it has been the differences which have been discussed, there is, in fact, considerable agreement on school objectives. Emphasis is given by most pupils to the instrumental functions of help with examinations, knowledge of jobs, and managing money. Expressive functions such as cultural development, outside visits, residential courses and education for leisure tend to be de-emphasized particularly among those who wish to stay on at school. It is those who wish to leave and those who are not considered by their teachers to have much chance of passing examinations who tend to rate these objectives more highly. In the prevailing examination dominated ethos, this may be a rationalization of lack of academic success.

Satisfactory comparison of the present findings with those of Morton-Williams is made difficult by differences in sampling, statistical analysis, and in the administration of the questionnaire. In 1966 when her study was carried out the statutory school leaving age was still 15, and the proportion of maintained schools in the country which were organized on comprehensive lines was only 8 per cent. With partial matching of respondents, however, some very tentative comparisons can be made on the basis of percentage response to the items in the questionnaire.

The overall impression is that schoolchildren in 1972 not only confirm the dominance of instrumental objectives in their view of school education but also interpret them more narrowly than the pupils in the Morton-Williams' research. The 1972 sample places even greater emphasis on examinations, which may reflect the growing popularity of the CSE examination. Good speech and writing, highly rated in Morton-Williams study, were given less importance by the present sample, as were social skills, studying a wide range of subjects and character training—with the exception of 'helping you to stand on your own feet'. The broader liberal aims of schooling seem to have declined even more in importance in recent years.

Within the pattern of similarities, differences between 'leavers' and 'stayers', and with social class, ability and sex have been described. The origin of these differences deserve further study. Is, for example, the greater importance which would-be leavers attach to expressive goals an acceptance of second-best to the real business of passing examinations? The present study indicates that for the high proportion of pupils adjudged to be incapable of passing public examinations this may well be the case.

There is no evidence that schools are becoming less examination dominated—indeed with the advent of CSE and the suggested school leavers' certificate the reverse may be true. (One of the advantages seen in raising the school leaving age is that pupils who would otherwise have left will now have the opportunity of sitting public examinations.) If, however, the extra year only serves to keep pupils in an environment in which their lack of academic competence is only too apparent, then it will not be the educational advance that has been claimed.

Appendix

How important do you think it is that school should:

1. Help you to become independent and able to stand on your own feet?
2. Help you to do as well as possible in examinations like GCE and CSE?
3. Help you to develop your personality and character?
4. Teach you things which will be of direct use to you in your job?
5. Help you to know about what is going on in the world nowadays?
6. Teach you plenty of subjects so that you can be interested in a lot of things?
7. Teach you things which will help you to get as good a job or career as possible?
8. Teach you about what is right and wrong?
9. Teach you to speak well and easily?
10. Teach you to be able to put things in writing easily?

How important do you think it is to:

11. Do drama in school, that is acting and reading plays?
12. Study poetry in school and read and learn poems?

How important do you think it is that school should:

13. Teach you things that will be useful in running a home, for example, about bringing up children, home repairs, decorating?
14. Teach you how to manage your money when you are earning and about things like rates and income tax?
15. Give you sex education?
16. Teach you how to behave so that you will be confident and at ease when you leave school?
17. Help you to make the most of yourself, for example, with your appearance?
18. Help you to learn how to get on with other people, for example, those you work with, your future wife or husband?
19. Give you interests and hobbies that you can do in your spare time?
20. Teach you about different sorts of jobs and careers so that you can decide what you want to do?
21. Help you to know what it will be like when you start work, for example, about hours and conditions?
22. Take you on visits to factories or offices or other work places to see the different sorts of jobs there are and what the work is like?
23. Take you on outings to places like art galleries, the theatre, museums or castles?
24. Take you on visits to places like the local fire station or town hall to learn what is going on in the world outside school?
25. Run clubs that you can go to out of school hours?
26. Take pupils away on holidays?
27. Arrange courses in which pupils live away from home for a while?

References

ALLEN, E. A. (1960). 'Attitudes of children and adolescents in school', *Educ. Res.*, 3, 65–80.

BARKER LUNN, J. C. (1969). 'The development of scales to measure junior school children's attitudes', *Brit. J. Educ. Psychol.*, 39, 64–71.

CARTER, M. P. (1962). *Home, School and Work.* Oxford: Pergamon.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1954). *Early Leaving*. London: HMSO.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1959). *15 to 18* (The Crowther Report). London: HMSO.

CLARK, A. S. (1967). 'The raising of the statutory school leaving age to 16 with special reference to the views of children of average or below average ability and to the views of their parents'. Unpublished MA Thesis, Univ. of London.

COMMITTEE ON HIGHER EDUCATION. (1963). *Higher Education*. (The Robbins Report). HMSO: London.

DEPARTMENT OF EDUCATION AND SCIENCE (1972). *Statistics of Education*, 1971, Vol. 1, Schools. London: HMSO.

FITT, A. B. (1956). 'An experimental study of children's attitudes to school in Auckland, N.Z.'. *Brit. J. Educ. Psychol.*, 26, 25–30.

HALSALL, R. (1971). 'A study of attitudes to school and to parents of high and low academic stream pupils in the fourth year of a comprehensive school.' Unpublished MA Thesis, Univ. of Keele.

HARGREAVES, D. H. (1967). *Social Relations in a Secondary School*. London: Routledge and Kegan Paul.

KNIVETON, B. H. (1969). 'An investigation of the attitudes of adolescents to aspects of their schooling'. *Brit. J. Educ. Psychol.*, 39, 78–81.

LOTWICK, W. R. (1965). 'Raising the school-leaving age —the views of school-leavers', *Univ. Coll. Swansea, Collegiate Faculty of Education Journal*, 17–20.

MORTON-WILLIAMS, R. and FINCH, S. (1968). *Young School Leavers. Schools Council Inquiry 1*. London: HMSO.

SHARPLES, D. (1969). 'Children's attitudes towards junior school activities', *Brit. J. Educ. Psychol.*, 39, 72–7.

WISENTHAL, M. (1965). 'Sex differences in attitude and attainment in junior schools', *Brit. J. Educ. Psychol.*, 35, 79–85.

Journal of Curriculum Studies **Collins**

In the November issue

One Project's Awkward Thinking about Objectives. *W. A. L. Blyth*

Curriculum Concepts and Conceptual Clarity. *D. K. Wheeler*

What can Teachers Contribute to Curriculum Development? *K. R. Gray*

The Politics of Political Education. *Guy Whitmarsh*

Curricula of Teachers' Colleges. *Philip H. Taylor and R. S. Adams*

Objectives in Geography Fieldwork. *David J. Boardman*

Nigerian Teachers' Reception of a New Syllabus. *P. A. I. Obanya*

It may be ordered direct from the Publishers

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada
Agathon Publications Services Inc.
150 Fifth Avenue
New York, N.Y. 10011

May and November
Annual Subscription £2.00 including postage

Interpreting Research: Sampling

Bruce Choppin, National Foundation for Educational Research

Introduction

Although the logic of research design may seem fairly obvious to those conversant with scientific methods it is not necessarily so straightforward for those who are not. While there are some who close their eyes to the possibility of any worthwhile research in education, there is probably still a tendency for many others to accept at face value any 'scientific' conclusion which appears in print. But papers reporting research need to be read with a great deal of care. This applies especially in the case of educational research, which, by its nature, is a difficult business in which compromises often have to be made. So, just as it is the responsibility of the researcher to report his methods, data, etc., with accuracy, it is the reader's task to approach the presentation and interpretation of research critically and with caution.

For example, one important factor affecting the outcome of educational research is that the validity of the conclusions of a piece of work is largely dependent on the way in which the project has been conducted and, in particular, the method of sampling used. The purpose of this paper, however, is not to present a treatise on the theory of statistical sampling. Neither is it intended to pillory particular researchers or research procedures. The aim is merely to point out the importance of sampling considerations and to illustrate the kinds of effect which can be caused by difficulties in sampling by examining three different reports of research.

Why Sample?

Sampling is usually carried out because it is uneconomic and impracticable to extend a research study to an entire population in which one is interested. It is thus a means to gathering indirect evidence on a general educational topic. Generalizations about the population are inferred from results obtained from the sample, but these are unlikely to be valid unless the sampling has been done correctly. Poor sampling can effectively destroy the value of a study almost before it is begun.

A good sample must be representative of its population, and this means that all relevant variations among the population must be reflected in the sample. Ideally, the elements will be chosen at random to eliminate deliberate bias. This means that every individual in the sample has a specifiable non-zero probability of being drawn in the sample, and that this probability is not affected by who else is selected. In practice, this is very rarely possible, and so compromises have to be made. One solution, often employed in national surveys, is to sample schools at random, but then to include all the children, or teachers, in the chosen schools. This is an example of a complex random sample. It differs from a simple sample, because the probability of a particular individual being included is dependent on whether others from the same area, the same school, etc., have also been included. Statistical tests and estimates have to be adjusted to allow for this. More frequently, samples are selected on the principle of convenience—the schools all lie in a particular town, or the teachers are all enrolled in a particular in-service course. Then randomness is missing, and it is more important to demonstrate the fairness of the sample, and to estimate and allow for any systematic bias.

On the other hand, a non-random sample *can* be quite good and may yield hardly less information about the population than a truly random one. The crucial consideration is that each sampling method should provide a basis for checking its fairness. For random samples this can be done by examining the variation in the sample itself. For other samples it is a matter of comparison with known population characteristics, and detailed arguing. An excellent exposition of this can be found in Kaplan's classic book on research techniques (Kaplan, 1964, p. 239, *et seq.*).

It will be clear from the above, that it will be hard to show the fairness of very small samples, especially where they are not random. In an extreme case, we cannot expect to learn much about all children from the detailed study of one child, or much about all secondary schools from research based in only one. If a generalization is to be credible, then the sample on which it is based must be reasonable. A good sample does not in itself guarantee a good research project, but a bad sample can largely cancel any value that a study might otherwise have had. Current researchers in education sometimes pay too little attention to this point.

We will illustrate these ideas by considering three articles all appearing in recent issues of *Educational Research*. They are neither outstandingly good nor bad examples of research reports. Each has been chosen for comment here because it demonstrates some particular facet of the sampling dilemma.

DOUGLAS, J. F.: A Study of Streaming at a Grammar School

(*Educational Research, Vol. 15, No. 2, February, 1973, pp. 140–143*)

Mr Douglas reports an analysis of data gathered from a boys' grammar school during the first eleven years of its existence. The study examines changes in patterns of examination performance as the school altered its policy from streamed classes to unstreamed classes.

Clearly it would have been difficult to carry out a study such as this on a random sample of grammar schools drawn on a national basis. Relatively few schools would have changed so much between 1960 and 1970 in regard to their policy over streaming. Mr Douglas' approach to the problem seems most practical. However, the limitations of using only one school must be noted. There is no reason for the reader of the paper to suppose that this school is in any way typical of grammar schools in general. In fact, its very newness suggests that it is not.

A second limitation deriving from the sample size of one, concerns the time factor. One might be inclined to surmise that the improvement in examination results during the decade might have been due to:

(a) changes in the teaching staff;
(b) developing relationships between the school and the parents; or even possibly
(c) increases in the size of the school library.

In fact any change in the school over the ten-year period might be thought responsible. In a larger study these alternative hypotheses would have been tested and perhaps rejected. With a sample of only one, an unlimited number of alternative but untestable explanations for the observed improvement could be made.

The final paragraph of conclusions can be quoted verbatim:

> In summary, this research indicates that less streaming of the grammar school ability band leads to social and academic gain for more pupils but with some loss in performance by the brighter boys. Perhaps too wide an ability range in the class would increase this loss intolerably.

This seems to refer to *grammar school* as a general educational phenomenon and not merely the one grammar school that had been the subject of study. As such the conclusion seems to be too sweeping. Although it provides useful indicators, this piece of research tells us very little about the effects of streaming at large. While it is interesting to read of developments at a particular school, it would be rash to generalize the findings to other schools.

CORTIS, G. A.: An Analysis of Some Differences Between Primary and Secondary Teachers

(*Educational Research, Vol. 15, No. 2, February, 1973, pp. 109–114.*)

Data were collected from a sample of teachers at the end of their second year of teaching, and these were analysed in conjunction with test and questionnaire results obtained during the final year of the college course. Some statistically significant differences between those students who became primary teachers and those who became secondary teachers were noted.

Dr Cortis gives few details about his sample, presumably to conserve space, and repeatedly refers us to his earlier work for more information on the study. He does explain at the outset that the paper is based on 127 primary teachers and 95 secondary teachers. This is not a huge sample, but is not as inadequate as was suggested in the press. Its representativeness is, however, in doubt. The author notes 'no differences in examination grades in educational theory and practical teaching' between 'respondents' and 'non-respondents', but that is no help unless there is evidence that the original group (respondents plus non-respondents) was itself representative. There are no details on educational background, age or even sex of the sample. Were they all men? Or all women? Did a higher proportion of women enter primary teaching—they certainly do in the population as a whole! A suspicious reader might wonder whether this effect alone might not account for the author's discoveries of personality difference between the two groups. Although commendable progress towards equality of treatment has been made in the teaching profession, it is a well established fact that men and women do not respond identically to personality tests, attitude tests, nor indeed on 'self report questionnaires'. Sex differences cannot be eliminated merely by ignoring them.

The final section begins:

> In summary, the primary teacher here appears a more 'rigid', insensitive type coming from a more densely populated area, with overtones of 'suspiciousness' and 'conventionality' distinguishing his behaviour. The secondary teacher, on the other hand, is more 'sensitive', more satisfied with his work and has more 'progressive' educational attitudes than his primary counterpart. He is less 'rigid', even inclined to 'laxity', and has more commitment to his academic main subject as a separate entity.

Taken solely as a comment on the 222 teachers involved in the study, this conclusion may be arguable, but is not objectionable. However, the weight of all the inferential statistics preceeding it shows that the author is concerned to make statements about a population not only about his sample. In the absence of alternatives, the population would seem to be British school teachers in general. Reading some of the other published material on this particular research reveals that in fact the sample was drawn entirely from the student body at one particular college of education a few years ago. This means that members of the sample were mostly young, and had little or no practical experience as teachers. They also had a special educational background not shared by the bulk of teachers nationally. On all these grounds, and particularly because they were only *prospective* teachers, they must be regarded as quite atypical of the teaching profession as a whole.

The paper closes with a paragraph, too long to repeat here, but well worth study, arguing that any shortcomings of the sample, are an inevitable result of the general difficulties involved in doing educational research. Although the point that 'purely random' samples are rarely the answer seems to be acceptable, this would not seem to explain the lack of interest in the composition of this sample portrayed in the earlier parts of the paper. One wishes that the author had tried to follow the suggestion that he quotes (Wall, 1968, p. 10) that 'the answer lies in the determination and estimation of sources of error'. Instead he appears to have ignored those aspects of his research design which might have been expected to produce bias in the results.

QUIGLEY, H.: The Pre-reading Vocabulary of Children Leaving Nursery School

(*Educational Research, Vol. 16, No. 1, November, 1973, pp .28–33*)

> A group of children in their last term of attendance at nursery schools were tested on their comprehension of words used in the first books of well-known reading schemes. No child recognized all the words and the vocabularies of individual children were seen to differ greatly. The implications of this finding for infant and nursery school teachers are discussed.

This neat study focusses on an issue of obvious importance, but as the conclusions reached are quite general, it is important to see whether the sampling procedures used were adequate. There are two types of sample to be considered here. Firstly, the sample of children that were given the test; and secondly, the sample of elementary reading texts that provided the words used in the test.

To take the second problem first, there appear to be few grounds for criticizing the selection of the reading schemes. A separate survey of infant school headmistresses revealed that nearly all children are faced with one of six reading schemes. These six were used as a basis for the test, and they themselves are listed at the end of the article. There seems little doubt, therefore, that the test contained words which were representative of the words met by children when they begin to read.

The sample of children to whom the test was given appears less than satisfactory, although the author should be commended for considering the matter, and for presenting the relevant details.

> This test was given to 50 children in their last term of attendance at four local authority nursery schools in South-East England.

This is a very small sample and might have been expected to lead to difficulties later. The social class distribution is checked against figures for the national population and an unexpectedly high proportion of working class children are found. This is noted but its possible implications for the results presented later do not seem to have been given much weight.

Section (b) of the results report 'no significant sex difference' when mean scores for boys and girls are considered separately. This is somewhat frustrating since it seems very unlikely that there would be *no* consistent differences in verbal development in the population as a whole at this age. The reported null finding seems to be merely an artifact of the very small sample employed. Similarly, it seems to have been impossible because of the small scale of the study, to show the expected relationship between word knowledge and social class.

Given these doubts about the social composition of the sample, and its inability to present evidence on anticipated relationships with sex and with social class one would hesitate to draw any sweeping conclusions from this research. In view of this the final section of the article would appear to have been somewhat overstated. The findings are tentative but certainly not conclusive, and the author might perhaps have stressed the need for further research on the problem.

Postscript

It would seem from this brief examination of three research reports that it can be very misleading to read only the last paragraph or last page of such a report. Of course, this would also apply to many other types of writing, but the point is particularly worth stressing here. Research reports tend, unfortunately, to be rather dry reading, and there is often a strong temptation to skip quickly to the end.

Each of the three papers commented on above represents a reputable inquiry into a problem worth studying. Yet each, it is contended, is weaker than it might have been because too little attention was paid to the choice of a sample.

It is usually the case that research in education provides evidence for propositions rather than any formal proof. In any study the results claimed have to be interpreted in the context in which the study was conducted—and that context includes the relevance of the sample. The responsibility remains with the reader to evaluate carefully the quality of the work, and weigh the evidence it produces, before accepting uncritically the conclusions of the researcher.

References

KAPLAN, A. (1964). *The Conduct of Inquiry*. San Francisco: Chandler.

WALL, W. D. (1968). 'The future of educational research', in WALL, W. D. and HUSEN, T. *Educational Research and Policy Making*. Slough: NFER.

The Utility of Taking Notes as an Aid to Learning

Michael J. A. Howe, Department of Psychology, University of Exeter

Summary

Does note-taking by students contribute to the acquisition of knowledge? Taking notes does provide a record of information, but since there are inexpensive alternative methods of reproducing knowledge, the extensive continued use of note-taking in education can be justified only if learning is facilitated thereby.

Much of the published research takes the form of broad comparisons of recall following experimental conditions in which learners took notes or made no notes, and the majority of such studies have produced no evidence for positive effects of note-taking. However, recent research has shown that in addition to recording information, note-taking influences learning by (a) ensuring that learners attend to information, (b) providing a version of the information that is particularly valuable to the individual learner, and through (c) direct effects of the activities involved in taking notes. Learning is strongly influenced by the particular note-taking strategy an individual adopts. It is suggested that the failure of the earlier comparison studies to produce firm evidence of the value of note-taking is due to their failure to exert precise control over learners' activities.

Introduction

It is widely agreed that the activities of individual learners are as important as are teachers' instructional methods in determining school achievement. However, comparatively little attention has been paid to assessing the effectiveness of the methods and techniques that students customarily use, or to ensuring that students receive training in making use of those techniques that have been shown to be of value. One very common practice is that of taking notes. In most forms of post-elementary education students write down what they consider to be important points from knowledge that is either presented orally by the teacher or is provided in printed form. The notes that a student makes give him a convenient record of information which is required. In addition, many teachers consider that the actual activity of note-taking has an educational value apart from the utility of the record which is thereby produced.

The fact that an educational practice is well-established provides no guarantee of its continuing effectiveness. In the case of note-taking, technological developments have led to increasingly inexpensive and convenient means of performing the function of reproducing information, traditionally provided by note-taking. Therefore it is important to know whether taking notes does have additional outcomes that influence learning. If taking notes has no such effects, it would be wise to restrict use of the practice to situations where no alternative means of recording required knowledge is at hand. However, if note-taking does influence learning, we would expect to find that techniques and strategies differ in effectiveness, and it would be valuable to discover which kinds of note-taking procedures make the largest contribution to performance in school tasks.

Additionally, if differences exist in the effectiveness of note-taking strategies, it is necessary to devise and evaluate training procedures by which learners can acquire the ability to make use of those strategies that are found to be the most efficient. Children at school have to learn how to make use of the methods and techniques that contribute to their acquisition of knowledge, and note-taking techniques form no exception.

The research that has been undertaken to investigate the value of note-taking procedures can be divided into two broad categories. Investigations in the first category have taken the form of attempts to compare the effectiveness for learning of note-taking and other procedures. The second category contains studies that have measured the precise outcomes of the activities involved in note-taking, and in subsequently making use of the notes. In most of the investigations the measure of learning takes the form of a score for recall of knowledge that has been made available in a lecture, rather than of scores measuring acquisition of rules, skills or concepts that are entirely novel to the learner. It is largely a matter of semantics whether the obtained measures of performance at recall tests are regarded as representing 'learning' or 'retention' of knowledge.

Comparison Studies

A number of experimental studies have compared the effects of learning procedures that incorporate note-taking with methods in which no note-taking is involved. Typically, scores of the recall of information following experimental conditions that require note-taking are either inferior to or not significantly different from scores following procedures that do not involve notes. For example, McClendon (1958) asked university students to listen to 14-minute lectures describing various aspects of communications. Three experimental groups were respectively asked to take no notes, detailed notes, or to take notes in their customary manner. The students were tested for retention of the information immediately afterwards, and again after a five-week interval. The findings showed that in neither the immediate nor the delayed test did the note-taking instructions have any effect at all upon the accuracy of retention. Eisner and Rohde (1959) required some of their subjects to take notes during a 30-minute lecture, and others were told to write notes on the lecture after it was finished. A 'True-False' recognition test was given immediately afterwards, and there was a further test three weeks later. On neither of the tests was there a significant difference in performance between the students who had taken notes during the lecture and after the lecture. Poppleton and Austwick (1964) measured the performance of grammar school students and post-graduate university students who learned statistics by one of two methods. Some of the subjects received the materials in a programmed instruction format, and the others read the same information in regular test form, but were encouraged to take notes. It was found that the two methods of instruction were equally effective for the post-graduates, but among the school children those who used programmed learning were more successful than the students who learned from materials in text form. In another investigation, described by McLeish (1968), adult subjects who listened to a lecture and took no notes at all were observed to perform better on immediate and delayed recall tests than individuals instructed to take either detailed or outline notes.

Peters (1972) suggested that students who are busy writing down points as they listen to a lecture do not hear other points. He predicted that note-taking would enhance learning at slower but not at faster rates of presentation. However, his experimental findings showed no significant difference between conditions in which presentation speed differed. In a study by Schultz and Di Vesta (1972) it was found that note-taking did not influence recall of statements, but that note-takers were less influenced than other subjects by certain kinds of organization in the materials. Ash and Carlton (1953) found that taking notes during the presentation of a film did not improve students' learning. This appears hardly surprising, since note-taking and film-viewing are not the most compatible activities. Taking notes might well have interfered with an individual's attending to the film.

In short, the few published experimental comparisons have produced no definite evidence that students retain more knowledge when they take notes than when they do not. However, it is quite likely that the lack of statistically significant differences is due to the fact that global comparisons of this kind do not provide adequate controls for a number of potentially important variables. An additional problem with the comparison studies lies in the implicit assumption that there is an identifiable, unitary, note-taking 'method' which can be compared with alternative educational methods. In fact, note-taking takes place in a range of situations wherein learners who differ considerably in age and ability take notes from information that comes in a variety of different forms. However, although the view has been expressed that only rarely can educational research give rise to widely generalizable statements (Glass, 1972), the majority of situations in which learners take notes have some aspects in common, so it is likely that experimental findings apply to a broad cross-section of situations that involve note-taking.

Yet another limitation of the experiments designed to investigate note-taking by making broad comparisons between conditions in which students do take notes and take no notes is that they have generally failed to consider the fact that in real life students normally make use of the notes they have prepared. It is important to assess the joint outcome of making and subsequently using notes, and this is not possible in the majority of published studies, since the usual procedure has been to remove notes from students upon completion. A study by Freyberg (1956) provides one exception. He found that students permitted to use their notes in order to revise information that had been presented in lecture form subsequently performed better on a test of retention than students who were unable to consult their notes. However, Freyberg's results provide no evidence that the students' own notes were any more useful for revision purposes than an alternative version of the information, such as a mimeographed summary that the teacher might provide.

Experimental research in the second of the two categories, investigating the direct outcomes of note-taking activities, has typically provided closer experimental control than has been available in the broad comparisons summarized above, and, perhaps as a result, the findings indicate that note-taking does indeed have discernible outcomes that have important implications for human learning. It is possible to list some precise consequences of note-taking activities. The most clearly apparent result of taking notes is to provide a record of verbal information. A second probable outcome is that taking notes may help a learner attend to knowledge he is attempting to assimilate. Thirdly, it is conceivable that in the process of reproducing information in notes the learner may encode the material into a form that he can remember more easily than the original version. A fourth possibility is that apart from the conceivably greater ease of remembering information that has been processed or encoded by the learner, the actual processing operations carried out by the individual as he takes notes may bring about increased retention of the knowledge.

An Analytical Approach

Thus there are at least four ways in which taking notes may help learners acquire knowledge. There is little doubt that note-taking does have the first of these functions, that of recording information which is presented in verbal form. The accuracy of such recording depends upon a number of factors, and in one experimental study the authors (Hartley and Cameron, 1967) noticed an alarming drop in accuracy of reproduction over the course of a 50-minute lecture. At the beginning the agreement between the content of notes made by students and a set of 'ideal notes' prepared by the lecturer was, on average, 70 per cent, but in the final ten-minute period this figure dropped to about 20 per cent. Similarly, Howe (1970a) found that even with a short (160 word) extract from a modern novel, university students, on average, were successful in recording the essential meaning of only eleven out of the twenty items in a passage judged by the experimenter to be important.

The second of the suggested functions of taking notes is that learners may be helped to attend to items they are trying to learn. Is there any evidence that note-taking has this effect? Research on learning in young children (Estes, 1970) has shown that in some circumstances large differences between normal and retarded children in performance at learning tasks, apparently reflecting basic differences in 'learning rate', disappear entirely when careful instructions are provided to ensure that all subjects actually attend to the same environmental events. There is no doubt that attention is a major variable affecting human learning. The available empirical evidence supports the suggestion that taking notes influences the direction of a student's attention. This is sometimes to the detriment of learning, as in the study by Ash and Carlton (1953) described earlier, in which students who had to take notes as they watched a film subsequently recalled less of the contents than control subjects, presumably because note-taking prevented them from giving full attention to the film. Howe (1970a) found that when students were required to take notes on a prose passage they later recalled ·34 of the items that they recorded accurately in the notes, compared with only ·05 of the items that judges observed were not recorded in subjects' notes. A number of separate factors may have contributed to this large difference, but it is likely that among them was the fact that learners had to attend selectively in order to make effective notes.

The third suggestion is that notes which have been prepared by a student constitute a version of the information which is more readily understandable and hence more valuable to him than an alternative version, for instance a set of notes prepared by the teacher. Two kinds of studies have provided indirect evidence on this point, showing that information which has undergone active processing and organizing by a learner is strikingly well remembered. First, research on organizational factors in human memory (reviewed by Howe, 1970b, and Postman, 1972) indicates that the ability to retain verbal information depends upon the learner being able to impose a degree of organization upon the retained items, even when the materials presented are selected randomly. When items are presented in a manner which clashes with an organizational scheme previously established by the learner, performance at learning tasks is poor, and the positive effects of repetition upon the learning of word lists are markedly reduced when subjects have to alter the manner in which items are organized (Tulving, 1966). Similarly, Bower, Lesgold and Tieman (1969) observed that subjects who were encouraged to impose different groupings on items in word lists from one trial to the next recalled considerably fewer items than individuals who were allowed to use the same grouping throughout.

The findings which have the greatest direct implications for note-taking situations are from experiments in which students were asked to reproduce information to which they had listened. Following the tradition of Bartlett (1932) and Allport and Postman (1947), who had demonstrated the importance of constructive activities carried out by the individual learner in the acquisition of knowledge, Kay (1955) and Howe (1970c) required adult subjects to make a number of separate attempts to recall meaningful prose materials to which they had listened. Both investigators found that individuals were strikingly accurate at recalling what they had themselves produced on earlier attempts at recall, both when the contents of such attempts were correct and when they were incorrect. This result emerged despite the fact that after each recall trial subjects listened to the correct version of the prose passage. Howe found that incorrect items appearing as part of a subject's attempts at recall were twice as likely to be repeated on a subsequent test as correct items that had been presented on as many as four separate occasions, as part of the repeatedly presented original version, but which had not previously been recalled. If an item was not correctly remembered on any of the first three trials, its probability of recall on Trial Four was less than 0·2, despite the fact that the learner had by then heard it no less than four times. On the other hand, items presented only once, but subsequently remembered on one occasion, had a 0·7 chance of being recalled again on the next trial. These results are based on highly reliable measures of retention of the meaningful contents of the narrative passage.

The importance of these findings for the practice of note-taking is that learners are shown to be far more accurate at remembering information which they have themselves reproduced than at recalling material to which they have merely listened. Note-taking provides one way of reproducing information that needs to be learned. Thus there is reason to expect learners to be very good at remembering the contents of their notes. Is this really so? The findings of Freyberg's (1956) study, mentioned above, confirm that notes can be useful for revision purposes, and Howe (1970d) has provided evidence that review sessions in which students read their own notes can facilitate recall. However, the crucial question is not whether or not the use of notes can contribute to learning, an assertion few would deny, but whether the notes that an individual has prepared are more effective than an alternative version, perhaps one prepared and distributed by the teacher, containing substantially identical information. We are currently planning research that will provide an answer to this question.

The fourth suggestion is that note-taking activities may have a direct influence upon learning, distinct from any effects of attending and from the outcome of making use of notes that have been prepared. One test of this suggestion is provided by measuring whether learning is reliably related to individual differences in the kind of notes students make, in circumstances whereby the notes are removed upon completion. There is considerable agreement that learning is related to the degree to which students' cognitive processes are involved in coding, integrating and transforming information that is made available (Bloom 1956; Ausubel, 1968; Howe, 1971, 1972) and we might expect that learners whose notes transform information they hear into 'their own words' would remember the contents more accurately than individuals who simply provide a verbatim reproduction of what they hear.

One of the findings of Howe's (1970a) experiment indicates that this is so. Judges examined the notes made by individual subjects, and measured both the number of words that appeared in a subject's notes, and the number of meaningful segments (out of twenty) from the original passage that were effectively communicated in the notes. (A correlation of over +0·9 between judges indicates that the scoring for meaningful content was highly reliable). Next, for each subject, the measure of the number of words used was divided by the score for meaningful content. The ratio thus obtained provides an indication of the 'efficiency' of note-taking; the greater the amount of content communicated, and the smaller the number of words required, the higher the score. The next step was to calculate the correlation between the score and the results of a later recall test. In fact there was a statistically significant correlation (+·53). This finding provides strong support for the suggestion that learning is influenced by individual differences in the kinds of note-taking activities, or strategies, that learners adopt.

Concluding Remarks

In summary, there is firm evidence for all four of the suggested outcomes of taking notes. Not only is a record of information produced thereby, but learning is facilitated by the effects of note-taking upon attention, by the use made of the individual's taking notes. In the case of the third effect, further research is required in order to determine whether an individual's own notes, once prepared, are likely to be more useful to him than notes distributed by a teacher, covering substantially the same material.

In view of these clear indications that note-taking has outcomes which influence learning, we might ask why the experimental comparisons described earlier produced no positive findings concerning the value of taking notes? The failure of the comparison studies to provide such evidence appears to be due to there being insufficient control over the actual activities of learners in the different experimental conditions. Learning is largely determined by what individual students actually do, and the latter may be controlled only to a limited effect by the experimenter's instructions. Thus, individuals who are instructed to take notes, but not given more specific guidance, are likely to vary considerably in their activities. Some students take hardly any notes, some take detailed notes, and individuals also differ in the kinds of notes they prepare. To study the effects of note-taking, it is necessary not only to examine the relationship between experimenters' somewhat vague instructions and subsequent performance at a task that measures learning as is typically the case in the studies comparing the effects of note-taking conditions with other methods but to investigate the relationship between what learners actually do in the note-taking situation and in other conditions, and their subsequent learning. It is probably for this reason that the experimental research investigating the four suggested outcomes of note-taking led to positive findings which did not emerge from the earlier studies. The crucial aspect of the more recent experiments is that they investigate the direct effects of what each individual actually does in the note-taking situation.

This point has broader applicability to research on the effectiveness of various teaching methods is likely that different methods of instructio influence learning only insofar as they encourage learners to adopt strategies that are more or less effective. The limited value of previous research upon the effects of teaching methods may thus be a reflection of the limited influence of instructional procedures upon learner's activities.

To return to note-taking, since it has been found that learning is greatly affected by the particular strategies adopted by individual students as they take notes (Howe, 1970a) it will be useful to direct further research to continued inquiry into the relative effectiveness of various strategies, and into development of procedures for giving children instruction and training in order that they will learn to make use of those strategies which have been found most effective. In the case of relatively mature learners, simple instructions may provide sufficient guidance for the adoption of effective procedures for the taking of notes. However, in the case of children at school, we need to provide ways of training individuals in the use of note-taking skills. Broadly speaking, it is well worth paying considerable attention to the development of methods which individuals can use to facilitate learning.

References

ALLPORT, G. W. and POSTMAN, L. (1947). *The Psychology of Rumors*. New York: Holt, Rinehart & Winston.

ASH, P. and CARLTON, B. J. (1953). 'The value of taking notes during film learning', *Brit. J. Educ. Psychol.*, **23**, 121–5.

AUSUBEL, D. P. (1968). *Educational Psychology: a Cognitive View*. New York: Holt, Rinehart & Winston.

BARTLETT, F. C. (1932). *Remembering*. London: Cambridge University Press.

BLOOM, B. S. *et al.* (1956). *Taxonomy of Educational Objectives. Handbook One: Cognitive Domain*. New York: David McKay.

BOWER, G. H., LESGOLD, A. M. and TIEMAN, D. (1969). 'Grouping operations in free recall', *Journal of Verbal Learning and Verbal Behaviour*, **8**, 481–93.

EISNER, S. and RODHE, K. (1950). 'Note-taking during or after the lecture'. *Journal of Educational Psychology*, **50**, 301–4.

ESTES, W. K. (1970). *Learning Theory and Mental Development*. New York: Academic Press.

FREYBERG, P. S. (1956). 'The effectiveness of note-taking', *Education for Teaching*, 17–24.

GLASS, G. V. (1972). 'The wisdom of scientific inquiry in education', *Journal of Research in Science Teaching*, **9**, 3–18.

HARTLEY, J. and CAMERON, A. (1967). 'Some observations on the efficiency of lecturing', *Educ. Rev.*, **20**, [illegible]–7.

HOWE, M. J. A. (1970a). 'Using students' notes to examine the role of the individual learner in acquiring meaningful subject matter', *J. Educ. Res.*, **64**, 61–3.

HOWE, M. J. A. (1970b). *Introduction to Human Memory: a Psychological Approach*. New York: Harper & Row.

HOWE, M. J. A. (1970c). 'Repeated presentation and recall of meaningful prose', *J. Educ. Psychol.*, **61**, 214–9.

HOWE, M. J. A. (1970d). 'Note-taking strategy, review, and long-term retention of verbal information', *J. Educ. Res.*, **63**, 100.

HOWE, M. J. A. (1971). 'Some comments on research into school learning', *Journal of School Psychology*, **9**, 51–4.

HOWE, M. J. A. (1972). *Understanding School Learning: A New Look At Educational Psychology*. New York: Harper & Row.

KAY, H. (1955). 'Learning and retaining verbal material', *Brit. J. Psychol.*, **44**, 81–100.

MCLEISH, J. (1968). 'The lecture method', Cambridge Monographs on Teaching Methods, Number One. Cambridge: Cambridge University Press.

MCCLENDON, P. (1958). 'An experimental study of the relationship between the note-taking practices and listening comprehension of college freshmen during expository lectures', *Speech Monographs*, **25**, 222–8.

PETERS, D. L. (1972). 'Effects of note-taking and rate of presentation on short-term objective test performance', *J. Educ. Psychol.*, **63**, 276–80.

POPPLETON, P. K. and AUSTWICS, K. (1964). 'A comparison of programmed learning and note-taking at two age levels', *Brit. J. Educ. Psychol.*, **34**, 43–50.

POSTMAN, L. (1972). 'A pragmatic view of organization theory', pp. 3–48, in TULVING, E. and DONALDSON, W. (Eds.) *Organization of Memory*. New York: Academic Press.

SCHULTZ, C. B. and DI VESTA, F. J. (1972). 'Effects of passage organization and note-taking on the selection of clustering strategies and on recall of textual materials', *J. Educ. Psychol.*, **63**, 244–52.

TULVING, E. (1966). 'Subjective organization and effects of repetition in multitrial free-recall learning', *Journal of Verbal Learning and Verbal Behaviour*, **5**, 193–7.

Book Reviews

Piaget in the Classroom
Edited by Milton Schwebel and Jane Raph
Routledge and Kegan Paul, 1974, 305 pp.
Hardback, £3·50 Paperback £1·50.

Piaget in the Classroom is a collection of essays by 11 educationists, most of whom, though apparently not the editors themselves, are former students of Piaget and who now lecture, or work as child psychologists, in North America. No one reading their book could doubt that Piagetians are on the side of the angels where infant schools are concerned. They applaud the Wendy houses, water-stands, sand trays and toys that can be found in every even half well-equipped infant room. Their difficulty is that although, as NFER research shows, almost all British primary school teachers have heard of Piaget, there is nothing essentially Piagetian about a sandtray as such. Russian infant classes look very like British infant classes and Piaget's theories have little influence in the Soviet Union. Some contributors are more aware of this than others. Hermina Sinclair, in particular, is concerned at the over-emphasis in our schools upon skill-learning at the expense of conceptual understanding. But even she doesn't fully face up to the powerful social constraints acting against individualized learning in schools. John Holt, who unlike most of these writers, has years of experience as a school teacher, has shown how many children will work through to a conceptual understanding of some apparently minor point in number so slowly that the parents and the principal start to ask for a few drill sessions to get the curriculum moving again. There is also a great deal of evidence from empirical research showing that teachers spend only a small part of their day on curriculum instruction. And we can reasonably suppose that a still smaller part is spent on work designed to develop conceptual thinking. We also know that in the classroom pupils are not all treated in the same way. Teachers are attached to, or concerned with, some children while they reject, or are indifferent to, others. Faced with all this it is not surprising that there are no Piagetian schools. For what distinguishes a Piagetian sandtray from a common-or-garden sandtray is no more than the use made of it by the teacher in that special questioning we know so well.

Perhaps the most valuable insight any teacher could get from this book is the realization that questioning children is a skill almost wholly dependent on *listening* to the childrens' answers. It is known from empirical research that teachers typically don't try to understand answers they think are incorrect, but simply rephrase the question or ask another child until the required answer is obtained. If reading Piaget could get this simple message across then he is worth reading on this point alone.

Some of the questions Piagetians put to children do seem odd. Who, in an ordinary frame of mind, would confront a child with six toy horses, two toy cows, and ask: 'Are there more horses here, or more animals?' That young children don't understand questions like this fails to impress me. Some of the writers do get beyond this level; Eleanor Duckworth in a passage most relevant to class teachers writes:

> Let's take a pre-reader, who is learning . . . his letters. He has happily learned the shape of a C, for instance, and draws it, for himself—but backward. 'No!' he is told, 'That's not a C! A C is like this.' An hour later, in another pre-reading exercise dealing with shapes, he is expected to realize that a square is still a square when it is sitting on its point looking like a diamond! How can he make sense out of all that?

This is the sort of careful observation that can lead teachers to a more sensitive understanding of childrens' learning difficulties.

The book helps to clear up two commonly held fallacies about Piaget's position: (i) that conceptual learning progresses by genetically determined stages and that, therefore, there is no sense in teaching before the appropriate stage is spontaneously reached and (ii) that since thinking and learning is internalized through activity, the provision of activity materials is, itself, sufficient. Hermina Sinclair argues that these propositions arise from a misunderstanding of Piaget's interactionist approach and do not take into account recent research.

One can't really recommend this book as an introduction to Piaget (it's too expensive for that anyway) but it should be useful to teachers who have already an interest in Piaget, but who are not quite clear how to translate this interest into classroom practice.

ROY NASH

Framework for Reading

by Joan Dean and Ruth Nichols

Evans Brothers, 1974, 95 pp., Hardback £2·80

Some of the books published in the last decade concerned with reading and language are not read by teachers but remain collecting dust, on the staff room shelf. Yet throughout many schools there is an air of discontent about reading standards: teachers seem desperate for help. Help of a kind not always found in books, for books too often offer only the theory. *Framework for Reading* is particularly welcome at this time because throughout its 100 pages theory is directly related to classroom practice and to individual children and their needs. It is a book essentially to be *used*, to help head teachers and staff develop a structured programme for reading. Where no such programme exists, teaching reading, however inspired and sensitive, can become so haphazard that difficulties are created not cured.

The introductory chapters concern language development, reading and motivation, but the major part of the book contains two detailed check lists to be used with children learning to read and those finding difficulty in achieving success. These checklists are also available as printed spirit masters for use on duplicating machines. Once the child's achievement and difficulties have been established the teacher can then turn to the pages which suggest specific materials and ways for helping children toward their various goals. The check lists are not intended to provide age norms by which to evaluate the child's progress vis-à-vis the general population but to seek information on what children already know and what they need to be taught. Such guidelines are essential to those teachers attempting to individualize reading within an environment rich in opportunity.

If criticism has been raised concerning reading standards in our modern primary schools, this book will surely enable teachers to build a framework for reading and so help to dispel any ideas that disciplined teaching cannot exist side by side with freedom and choice.

In order to reduce the price and therefore reach the student population it is hoped that this book will be made available in soft covers.

BETTY ROOT

The Disabled Schoolchild

by Elizabeth M. Anderson

Methuen, 1973, 377 pp., Hardback £5·00,
Paperback £2·40

Although there has been an increasing belief among those professionally involved that physically handicapped children should be educated in ordinary rather than special schools wherever possible, little evidence has been available either to support or challenge this view. On official figures, 10,200 of the 22,257 physically handicapped school-children in England in January 1971 were attending ordinary schools. How do the academic achievements of these children compare with those of their non-handicapped classmates? Does attendance in ordinary schools cause emotional strain in physically handicapped children? Do these children suffer from social isolation and teasing in the ordinary class? Are their needs for physiotherapy or speech therapy neglected? Or is it that only the mildly handicapped are placed in ordinary schools anyway? What special provisions are made for physically handicapped children in ordinary schools, and could better provisions alleviate existing difficulties or enable greater numbers of physically handicapped children to attend these schools? These are the kinds of questions on which evidence is needed. In her book Elizabeth Anderson answers many of the questions. She bases the answers primarily on her own very careful study of 99 physically handicapped children attending ordinary primary schools in the south of England, which forms the core of the book. She also uses information obtained from visits to ordinary primary schools which are making special arrangements for physically handicapped children, both in this country and in Scandinavia, where there is an active policy of integration.

The detailed study made of the social adjustment and academic attainments of the junior school children with purely physical handicaps (predominantly orthopaedic) shows that they differed very little from their non-handicapped classmates on most of the measures. These children are obviously satisfactorily placed in ordinary schools. The position is less clear for the physically handicapped children with neurological abnormalities (mainly cerebral palsy and spina bifida) in the study. These children were experiencing more difficulties than the controls both in their schoolwork and overall adjustment. Miss Anderson warns against jumping to the conclusion that the physically handicapped children with neurological abnormalities are misplaced in the ordinary school. Although three-quarters of the children were poor attainers school performance was adequate when related to intelligence. Other recent research shows that attainments in special schools are also poor and it is unlikely that these children would have made better progress in such schools. It is clear that many physically handicapped children with neurological abnormalities need special educational help. Miss Anderson takes the view that this can be satisfactorily provided in ordinary schools. She suggests a number of possible arrangements for this provision of help, such as special classes or a 'clinic' to which the child can go for specialized help when necessary. Not everyone would agree that the ordinary school is the best place to provide this special educational help. Comparative studies evaluating the success of special schools and of various schemes providing special help in ordinary schools are necessary to resolve this question.

Study findings do support the belief that many children at present in special schools for the physically handicapped could be satisfactorily educated in ordinary schools. Over a half of the children in the study sample had actually been considered for special school placement at some time. These children were similar in their academic attainments and social adjustment to those for whom special school placement had never been considered. The severely handicapped children, who constituted one-fifth of the sample, also did not differ in educational attainments or social acceptance from the less severely handicapped children.

At present there is often hesitation about recommending physically handicapped children for ordinary school placement because of lack of special physical facilities, provision of therapy and special educational help. The final section of the book provides, for the first time, an extensive discussion of the types of special arrangements that are being made in ordinary schools both in this country and Scandinavia, of their inadequacies and strengths, and of possible improvements. Many severely handicapped children are being successfully placed individually in ordinary classes. Miss Anderson suggests that several of the inadequacies which often occur in this arrangement—lack of modification of buildings or provision of special equipment largely due to ignorance of the children's needs, insufficient information to teachers about the child's physical management and educational capabilities, deficiencies in provision of extra teaching—could be overcome by the appointment of an adviser in special education with particular responsibility for physically handicapped children in ordinary schools. Some authorities are starting to make such appointments. Where there are substantial numbers of physically handicapped children who need special physical facilities, or special educational help, individual placement in ordinary schools may not be feasible. There are two other main possibilities. In one arrangement which is being tried out in a few areas in this country, and extensively in Scandinavia, physically handicapped children are placed in special classes in one particular primary school in the area. This system probably could be useful for children needing special educational help but results in no real integration unless there is considerable forethought by those responsible for planning the unit and by the teachers themselves. In the other type of arrangement special adaptations, provision of ancillary help, transport and therapy are made in one particular primary school but the children are placed in ordinary classes. It is suggested that this is a particularly promising approach. Expansion of these various types of arrangements would allow many children at present in special schools to enjoy the advantages of ordinary schooling without losing those of special schooling.

Integration into ordinary schools is an issue concerning all handicapped children, not simply the physically handicapped. This book will be invaluable to anyone involved in special education.

DORIA PILLING

Compensatory Education for Children, Ages 2 to 8

edited by Julian C. Stanley

The Johns Hopkins University Press, 1973, 212 pp.
Paperback £1·40

> In my opinion, the educational system of this country needs to be reorganized.
>
> I have personally come to the conclusion that compensatory education as presently understood, is counter-productive. It seems to stem . . .from . . . prejudice, do-goodness, conscience-salving

and maintenance of the status quo and . . . from reluctance to tackle the larger project of revamping our entire educational structure.

Thus begins and ends this book. Both Julian Stanley in his introduction and Scarvia Anderson in her final critique would place more faith in restructuring the whole public school system than in early intervention schemes to improve the educability of children.

The other five contributors reveal similar doubts but are more defensive, perhaps because they are describing and assessing their own experiments. Some are self-contradictory and some mutually contradictory either in fact or interpretation, or to the quasi-scientific spirit of the papers. For instance, Sigel states: 'Other treatment programmes are possible, e.g. change the economic system . . . or attack insidious racism . . .' but 'these treatments do not engender the same respect as pre-school education.' (Respect on the part of whom?) He offers no evidence to support this conclusion whereas elsewhere in the book far less sweeping statements are supported by sophisticated evaluation techniques and batteries of tests (e.g. that children of middle class parents learn from watching Sesame Street). 'Thus', Sigel goes on, 'our society has made its choice . . . by assigning relatively large sums to pre-school systems'. This hardly matches the statement made by one ghetto resident (quoted by Shipman): 'We're sick and tired of graduates with clipboards.' (One wonders: Who are 'our society'? How were 'they' given an opportunity to choose?) Shipman continues: 'the . . . reasons for discontent and disenchantment are the stark and undeniable facts that the people living in the ghetto—the "subjects" of so much recent investigation—have rarely received any visible benefits from all the research flurry or had any control over what research was done or how it was to be done.'

Ball and Bogatz (in their paper on Sesame Street—of special interest to readers in England) confirm this view when discussing the Peabody Picture Vocabulary Test. It includes only two pictures of blacks—a spear carrier and a porter. The point that they are making is that on this test there is, in IQ terms, a 10-point difference favouring the whites. They discovered that blacks were suspicious: 'We were . . . asked by black leaders whether we were simply trying to prove that their children were dumb.' They go on to affirm that they intended no racial comparisons and presumably reassured their questioners upon this point. One cannot help wondering whether, given a choice, their interlocuters may not have opted for changing the economic system or attacking insidious racism in preference to pre-school intervention schemes.

But it would appear that such programmes have other benefits, not perhaps originally intended. Shipman sums up the value of her programme: 'First' (and note that she puts this first) 'there has been immediate pay-off in the communities involved in terms of dollars spent in hiring local employees and purchasing necessary test supplies to say nothing of giving research workers from Princeton University the opportunity to have 'the interesting experience of meeting poverty groups'.

It would be all too easy to be cynical about this book, but this would be unfairly destructive. The contributions are all presented clearly and simply so that anyone can read and understand them (especially if you begin at the back with Anderson's useful clarification of the terms 'compensatory', 'intervention', 'remedial', 'special') There is, in fact, no attempt to blind with science or obfuscate with esoteric jargon. It is a valuable and stimulating addition to the field. It does perhaps, unintentionally, give more ammunition to the 'education cannot compensate for society' school than it does to the committed intervenors, if they exist. For those interested in research it raises controversial issues around measuring instruments, evaluation techniques and the validity of testing. The choices made of what conclusions are or are not substantiated with statistical evidence are also worthy of thoughtful consideration.

Perhaps the last word must go to the 61 per cent of 'lower socioeconomic' mothers reported by Shipman as feeling that political candidates run more for personal advancement than for accomplishing campaign promises. Their native wit does not seem to have been dulled by the lack of early intervention.

SHIRLEY HADI

Teaching in the Urban Community School

by Norman Garner

Ward Lock Educational, 1973, 118 pp.,
Hardback £2·00, Paperback £1·20

Norman Garner was head of education at the Ethel Wormald College of Education until his untimely death, and was well known for his work in the City of Liverpool, especially in connection with the linking of colleges with EPA schools. At his death he left the synopsis for this book, which has been developed by Eric Midwinter and other friends and associates of Norman Garner.

The resulting book is a most important work for teachers in EPA areas, since it sets out to examine the needs of children in these areas with a refreshing and unbiased eye. But it is also of the utmost value to teachers in urban schools generally, since it emphasises the problems of modern urban life and the preparation that will be necessary to prepare children for the cities of the future, and the many problems as yet unknown which will face them as adults. It rightly emphasises the development of character and of human concepts which influence the growth of a child. Of the many interesting chapters, I would particularly recommend Jim Parton's account of social conditions in modern urban areas and Norman Garner's own comments on the development of the urban child and the urban curriculum. Paul Harman contributes a very penetrating chapter on discrimination from his point of view as Director of the Priority Everyman Community Theatre in Liverpool and Eric Midwinter makes his usual searching comments on the urban community and the urban school. There is much to be learned from this book and it is a fitting memorial to the work of Norman Garner.

Frank Plimmer

Pre-school and Primary Maths

by Ernest Choat

Ward Lock Educational, 1973, 107 pp.,
Paperback £1·35, Cloth Edition £2·50

The last decade has witnessed a revolution in the organization and the curriculum of the primary school, especially with regard to the character of mathematics teaching. The next decade may well see further great changes, particularly in the area of pre-school provision. Accompanying any expansion in the number of places for under-fives will arise the question, which mathematical abilities are the children developing at this age, and how can the nursery best promote them?

The collection of papers in this book deals with mathematics learning and teaching in both pre-school and primary school. Most of the papers were written for a working party which formed part of the Second International Congress of Mathematical Education held at the University of Exeter in 1972. A glance at the chapter headings suggests a comprehensive range of topics and perspectives, including the approaches taken to the mathematics curriculum, teaching, assessment, and teacher training by British nursery, infant and junior school teachers; as well as by teacher trainers, educational researchers and academics from Japan, USA, Hungary and France. Nevertheless, the contributors share a common purpose, and although the details of their accounts reflect their particular interests, there is a common set of problems which each of them faces and which appears over and over again in different guises.

One such problem arises out of the necessity of defining the mathematics curriculum relevant to children of different ages. The content of the subject must be reconciled with the stages of development of children's thinking. Several of the contributors question the assumption that on entry to primary school children have acquired a basic numeracy, pointing out that the ability to count by rote must not be mistaken for an understanding of the significance of the number-word series. The problem then is to specify these pre-mathematical abilities and to suggest how the nursery school might ensure that children have acquired them, especially in those cases where their home environment lacks experiences appropriate to their development. Amongst the abilities suggested in these papers are familiarity with the character of the natural world and knowledge of the vocabulary and concepts relevant to a mathematical description, e.g. relational terms such as 'bigger', 'smaller', 'more', 'fewer', 'faster'. 'slower', etc; also included are the basic skills of analysis of features, classifying and grouping.

Another problem confronted by all teachers concerns the conditions under which the child's mathematical abilities are best promoted. There is a good deal of agreement amongst the contributors to the present volume that, following Piagetian theory, the best approach to teaching abstract concepts in mathematics is by providing children with concrete experiences taken from everyday life. Many imaginative examples of the approach proliferate throughout these papers.

The inadequacies of both highly structured learning by drill, and laissez-faire child-centredness are pointed out. If there is any consensus of approach amongst these authors it is, in the words of one of the authors, 'child-centred but teacher controlled'.

However, the account by a junior school head of difficulties encountered in introducing new methods to the primary school is instructive. Irrespective of the benefits such methods may offer for improving the quality of mathematics teaching and learning, the strain that their introduction may put on the traditionally-trained teacher is not to be underestimated.

Much of the uncertainty and difficulty which is being experienced by teachers may well stem from the speed with which the revolution in mathematics teaching has taken place. The problem is one of managing innovation in such a way that teachers can readily assimilate new methods and apply them with confidence.

Two attempts to co-ordinate the introduction of new methods are described. At the local level, the Lower Valley project was an attempt to plan and prepare teachers for innovation through meetings and courses in local centres; and at the national level, research and development projects were initiated to help provide practical guidelines and resources for teachers of primary mathematics.

Anyone looking for a systematic account of modern methods in pre-school and primary mathematics teaching will be disappointed by this book. Although many ideas and practical suggestions may be culled from these papers, their principal merit is in illustrating some current thinking on mathematics and highlighting the difficulties encountered in introducing new methods into the curriculum, many of which are not specific to mathematics but are equally applicable to all aspects of the pre-school and primary curriculum.

MARTIN WOODHEAD

Family History in Schools
by D. J. Steel and L. Taylor
Phillimore, 1973, 180 pp., Paperback £1·50

History with Juniors
by Michael Pollard
Evans Brothers, 1973, 71 pp., Hardback £1·10

History teachers have been on the defensive for some years. Pressure for change in content, from the separate lobbies of world history and local history gave way to an emphasis on method: 'It doesn't matter what you teach so long as children acquire the skills and methodology of the historian'. No consensus has emerged, and history teachers, divided among themselves, have lost ground in the struggle for space on the time table.

Two principles lie at the root of much recent curriculum change: that children should begin by studying topics close to their own experience; and Bruner's thesis that children can acquire the essential structure of any subject as long as it is taught in an intellectually honest way. Family history may claim to satisfy both these criteria, and give teachers and children scope for an inquiry that involves them in investigating primary evidence, unique to each child, on an area in which he has the satisfaction of being the authority in his own field of research.

Messrs Steel and Taylor make out a good case for family history as the root of all other forms of historical enquiry. In primitive societies the only history was the collective memory of the ancestors. Today it could be the missing link in the chain of biography, local and national history. A child's memory of his own life is his first primary source; the memories of parents and grandparents provide oral evidence of the past closest to him, the connection between himself and history. Starting with the comprehensible unit of his own life span, he can acquire a sense of time, continuity, and sequence, and by investigating the experiences and changes in life style of older members of the family, he will understand one of the basic historical concepts—change.

Family History in Schools is based on the experiences of 18 teachers working with 700 children in primary and secondary schools in Berkshire and Hampshire. They tried out variations of a clearly structured scheme in which historical concepts and techniques were introduced in a progressive order of difficulty, in three phases: Who Am I; Involving the Parents; The Historian At Work. Transcripts of taped interviews; reminiscences of parents; family rent books, greeting cards, identity cards and gas masks, photographs, flat irons, apprentice indentures and documents of all kinds contributed to the range of sources the children handled. They acquired skills of labelling, sorting and card indexing, and when the primary evidence was exhausted, explored the more public sources of magazines, newspapers, trade directories and advertisements to fill out the picture of social conditions.

This is a practical, if discursive and disorganized book. The authors recommend respecting skeletons and leaving them in the cupboard, but seem a little ingenuous in suggesting that a child with 'special family difficulties' could be satisfied by studying the teachers' family, working on other children's evidence, or organizing the card index. The idea of collecting information from local old people is more promising. The authors append a set of teachers' reports which honestly admit the difficulties, both in guiding individual children to order the wealth of family material, and in the dearth of back-up resources on recent social history.

Michael Pollard's *History with Juniors* is a slighter, more impressionistic and anecdotal account, drawn mainly from his own teaching experience and touching on local history, oral traditions and folk lore as well as on family history. He too believes that children should begin with the area of the past closest to them, and that recent history based on primary sources offers them the challenge of using the historian's methodology, of appreciating sequence and contemporaneity, and understanding reasons for change. His message is crisply stated: the demolition of the Chiswick Empire has more relevance than the place where Caesar may have crossed the Thames. But the teacher in downtown city areas, expecting more detailed guidance on how to investigate signs of urban change will be disappointed, for Mr. Pollard never pursues the fate of the Chiswick Empire, but is diverted to muddy field work in Dorset and traditional country customs—after a more useful piece on the industrial development of Swindon.

Teachers in urban areas who need most help will not be satisfied with either of these books. But both go some way to providing teachers with a rationale for developing a method of studying aspects of the past closest to the hearts of children.

MARY WORRALL

The Nature of Human Values

by Milton Rokeach

The Free Press, New York, and Collier Macmillan, 1973, 438 pp., Hardback £6·95

For more than twenty-five years Milton Rokeach has been engaged in the study of belief systems. In the *Open and Closed Mind* he was primarily concerned with the acceptance and rejection o people or ideas; in *Beliefs, Attitudes and Value* he argued cogently for making values the core construct of the social sciences; and now in *The Nature of Human Values* he presents us with a wealth of empirical information about the distribution of values in American society, and the relationships of values to attitudes and behaviour.

A value for Rokeach is 'an enduring belief that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or end-state of existence'. He thus distinguishes 'instrumental values' concerned with 'modes of conduct' and 'terminal values' concerned with 'end-states of existence'. These conceptions are operationalized in the Value Survey in which a respondent is presented with a list of 18 terminal values (e.g. 'a world at peace', 'freedom') and 18 instrumental values (e.g. 'honest', 'helpful') and asked 'to arrange them in order of importance for YOU, as guiding principles in YOUR life'. The values, each accompanied by a short descriptive phrase, are printed on small removable gummed labels so that the respondent is free to go on re-ordering them until he is satisfied with the end result.

The values of American society are apparently fairly homogeneous with, among the terminal values, 'a world at peace', 'family security' and 'freedom' valued highly, and 'social recognition', 'pleasure' and 'an exciting life' somewhat less important. However, there were a number of differences between the various sub-groups. For example, the poor differed from the rich in placing high value on 'clean' and 'a comfortable life', and in other ways which are interpreted as giving some support for the notion of differing cultures of poverty and affluence. A number of differences were also found between black and white Americans, but, when income and education were controlled for, most, with the exception of the importance attached to 'equality', disappeared. Rokeach thus rejects Moynihan's (1965) view that there is a self-perpetuating black culture of poverty.

From his studies of values and attitudes, and behaviour, Rokeach develops a two-value model of political ideology, based on the relative valuing of 'equality' and 'freedom', which he supports by content analyses of the works of Lenin, Hitler, Goldwater and various socialists. In experimental studies, he is able to show that by bringing into the open inconsistencies between a person's self-conceptions and values it was possible to bring about changes in certain 'target' values which persisted for more than 15 months.

All this is very impressive, but it makes one scrutinize the Value Survey very carefully to see if it will bear all the weight that Rokeach tries to put on it. And here serious doubts arise. Do values form the strict hierarchy that is assumed? Aren't people really being asked to choose between answers to distinctly different questions, for example, 'what sort of person do I want to be', 'what sort of world do I want to live in'? Does everyone have the same amount of 'value energy' as is implied? Can people accurately rank 18 items (most of the differences arise in the middle range)? High and low ranks appear to have different meanings for different people, for example, some seem to be giving low ranks to values that are important but taken for granted. There appear to be important omissions from the lists of values—truth and health, for example, are not included among the terminal values. And I could go on, but I hope I have said enough to show my misgivings about the basic measuring instrument.

The Nature of Human Values is then a curious mixture. On the one hand, it confirms Rokeach as a researcher of powerful insight with a sound sense of the important and a flair for giving the lead to others. The two-value model of political ideology and the theory of cognitive and behavioural change are major advances. However, the Value Survey on which so much rests seems decidedly shaky, and it tempts Rokeach into vacuous and inflated assertions.

Rokeach, himself, draws attention to Kaplan's law of the instrument: 'give a small boy a hammer and he will find that everything he encounters needs pounding'. It is to be hoped that there will not be a rush to use the Value Survey simply because it is there.

ALAN SMITHERS

Reason and Compassion

by Richard S. Peters

Routledge and Kegan Paul, 1973, 128 pp.,
Hardback £1·80

Anyone looking for a new or radical approach to problems in moral education under such a global title will be disappointed. Rather in seeking to present a moral position based on the use of reason tempered by the passions, Professor Peters acts, in this collection, largely as an eclectic exponent of some of the most influential theories in moral learning and development. This slim volume comprises three lectures given as the Lindsay Memorial Lectures at Keele in 1971 and a fourth delivered to the Society of Friends.

The first lecture starts from the scarcely disputable premise that 'morality is complex'. It is more than a traditional authoritarian code at one extreme or, at the other, a number of individual subjective principles which a man can decide for himself. For, Peters maintains, above some basic social rules, such as non-injury or care of the young, there must be certain principles of fairness, consideration of others' interests, and so on, which provide criteria for what is to count as a reason in ethical questions. Referring to modern ethical theory the distinction is thus early made between the invariable *form* of moral consciousness and the *content* of particular moralities.

This is taken up in the second lecture, 'Morality and Moral Development', where a parallel is drawn with theories of the developmental psychologists about the stages of growth of the child's moral consciousness. Favouring Kohlberg's refinement of Piaget and with supplementary evidence from Freudian and social psychology, Professor Peters adroitly expounds the now well-known view of the child who begins 'egocentrically' by seeing rules as connected with punishment and rewards, later, 'transcendentally' backed by the praise or blame of his parents and friends and, finally, he passes, from this heteronomous morality to a more autonomous position where morality is seen as alterable and depending on consent and reciprocity. Autonomy is marked by three features: genuineness, reflective criticism and strength of will.

The more crucial and interesting questions for educators are now opened up in discussion. What sort of 'concrete' experiences should the teacher provide to enable the child to grasp principles unifying various experiences? What sort of content should pupils be exposed to in literature, history, religion and social sciences in order to enlarge their scope for making more informed and imaginative choices in their own conduct? How, more specifically, is it possible to sensitize the child to a consideration of other peoples' interests? Professor Peters steers a mid-course between the arbitrary fiat of the headmaster and progressive schools which 'insist *from the start* on children learning only what interests them, on making their own decisions and running their own affairs' and advocates a firm structure within which pupils are allowed increasing discretion and participation. Guidance to avoid bewilderment is obviously needed, yet how to do this without incurring the charge of indoctrination (even as tautologically defined as on p. 71), or when the transition should be made to exhorting the child to discuss, think through, and assume principles for himself, is not really adequately explored. Either way the teacher (and the parent) is in a difficult and responsible position.

In the third lecture Professor Peters contends Hume was wrong to separate reason and the passions. Instead he prefers to call objectivity, impartiality and such like the 'rational passions' and shows how in the form of emotions and motives they are connected, not just more obviously with feelings, but with our cognition or ways of understanding situations. The contrast should rather be characterized by different levels of life from the 'irrational', through the 'unreasonable' to the highest 'autonomous' level, which although not governed solely by reason is critical and self regulating.

At first sight, after the evolutionary progress of this chapter, Lecture Four, which deals with the *religious* dimension of a rational morality, seems somewhat antithetical, especially in view of the concern by Professor Peters *inter alia*, to distinguish religious from moral education. Closer inspection reassures that this is written from a non-doctrinaire standpoint, and everyday illustrations reveal how 'religious experience enables us to understand the facts of a more mundane experience in a new light'. Even if in the end Professor Peters is rather pessimistic about the ultimate pointlessness of life there is some indication of how life as lived can be viewed in a meaningful perspective and with compassion.

Despite the claims of the dust jacket, and the lucid though sober exposition, this book might prove solid reading for those without some philosophical training or familiarity with present thoughts in moral education. Few concessions are made to the layman especially where Professor Peters excels in his criticism and application of the Utilitarian and Kantian systems, Piaget and Kohlberg, but it is instructive to have the present views of one of the foremost philosophers in this field set out in such a concise and masterful way.

MONICA TAYLOR

Organisation and Memory

by Peter Herriot, Josephine M. Green and Roy McConkey

Methuen, 1973, 117 pp., Paperback £1·25

This book is a pleasure to read because it adopts the unfortunately rare strategy of underwriting rather than overwriting. The authors also dominate the material and the results and know exactly where they have said anything. In consequence the review and the report of experiments are compact, concise and neatly cross referred. The subject matter also has a certain novelty in the field because as the authors remark, few have tackled the subject of long-term memory in the subnormal and it is also true that fewer still have looked at the organization of memory in this field. Most of the studies quoted are drawn from the school centering round H. H. Spitz at Bordentown, New Jersey and reports of the work of I. R. Gerjuoy are central. As one might expect at this time the studies of organization in memory are heavily dependent on the work of Mandler, Postman and of course Tulving.

There are 11 short chapters, the first three of which are devoted to the definition of terms such as organization, process, cue and coding, the possible effect of organization on recall strategies, the role of organization in input-coding and retrieval and the currently absorbing topic of cues and their capacity to evoke items from store in unexpected quantity. The next two chapters discuss the possibility of the development of memory organization in children and a brief review of some work in this field, mainly concerning 'clustering' in subnormals. The sixth chapter sets the stage for the following reports of what would appear to be 12 experiments. The conditions surrounding these experiments are given in impressive but over concise summary in Table 1 on page 56. The remaining pages are devoted to reports of the experiments which occupy Chapters 7, 8, 9 and 10. Chapter 11 is a summary of conclusions. Measures of organization are described in Appendix 1 and a test of basic recall ability in Appendix 3.

The experiments themselves are concerned with the relationship between clustering (or organizing) and the capacity to recall. These are found to be only loosely related in the subnormal. Other experiments are concerned with the effects of cueing and capacity to aid recall if it occurs on input or at retrieval. Cues at retrieval may aid clustering but may not necessarily increase the amount recalled. The only form of assistance which increased amount recalled over the control condition in this instance was that in which both presentation and recall cues were category labels. Summaries of results of this work occur on pp. 71, 74 and 84.

The authors conclude that recall can be aided by process interventions such as category cueing, active sorting and general instructions. They believe that they have given evidence of secondary processes in the subnormal at least where Verbal mental age is around seven or eight years. The capacity of the individual subnormal to select and adopt an appropriate strategy, even when developed for use, is not necessarily strongly established by these experiments say the authors in the last chapter (p. 99).

This very brief book therefore represents a limited but perceptible step forward in the study of the subnormal and whatever it may amount to ultimately, seems to this reviewer to deserve commendation as a sincere, workmanlike and interesting attempt to apply current memory theory to the illumination of cognitive deficit in the handicapped. It uses tried experimental method and is simple, consecutive and direct. Naturally the reader will not expect a faultless story. The main weakness is the book's merit is its Baconian style. Some of the descriptions of experiments in particular are so short as to be inscrutable without resort to the original reports in journals. In some respects therefore, it remains a book for the expert although its basic simplicity deserves a wider readership. It might also be said to overlook long established evidence for the capacity of the subnormal to organise, although in justice to the modesty of the reviewer no sources will be quoted here. However, in general it is one of the better buys in books which today threaten to eat authors (and readers) when the reverse should be true.

N. O'CONNOR

The Open University Opens
edited by Jeremy Tunstall
Routledge and Kegan Paul, 1974, 191 pp.,
Hardback £2·50, Paperback £1·00

This book is a collection of articles by education experts, Open University staff and students. It is edited by Jeremy Tunstall who was senior lecturer in sociology at the OU and is now Professor of Sociology at the City University, London. The book takes a critical look at the first three years of the OU. Brian MacArthur suggests that in years to come those who have said that history will judge the National Health Service and the OU as Britain's major achievements of the second half of the twentieth century may be correct. The OU has many unusual characteristics: its academic year starts in January, its students who are spread throughout the United Kingdom require no formal qualifications for admission. They are mostly adults who study on a part-time basis using correspondence materials and radio and television broadcasts for the major part of their instruction. Although the OU employs lecturers many of them never deliver a lecture. There is no student campus for this university where the administration is largely done through a computer to which the students are known by number. The uncertainty about the place in higher education of the OU and the status of its degree is clearly brought out. However, the institution has already weathered some severe crises. It is thought that 'a paper proposing the abolition of the University was on the desk of Iain MacLeod the night he died'. The Conservative Government endorsed the university although it cut the budget by £1m over three years. The postal strike also severely affected the University.

The book sets out to be an 'interim and completely unofficial report' on the OU. It is frankly critical of many aspects of the organization. One of the most serious problems of the OU are its 'casualties'. In the first three academic years about 30,000 students have failed and/or dropped out (some may have dropped out temporarily). About half of all students who start the Science and Mathematics foundation (first-year) courses each year drop out. Leslie Wagner, lecturer in economics at the OU, points out that the drop out rate at the OU would have to rise by about 85 per cent for the cost per graduate to be the same as at a traditional university. There are several personal articles by students which make interesting reading as they reveal the diversity of the student population and the motivation behind their doing an OU degree. On the subject of students some other interesting facts that emerge are that although the student population is largely drawn from the middle class their parents were working class. It would also appear that women students last the course better and have a higher success rate than the men.

This book is critical of the OU but although unofficial, it is a useful collection of information and opinion on what must surely be one of British education's most innovatory institutions. It will be of particular interest to present and prospective students and all those interested in the OU. It will also appeal to overseas educationalists who may be considering undertaking a similar venture in their own countries. JANET MAY

Student Counselling in Practice
by Audrey Newsome, Brian J. Thorne and Keith L. Wyld
University of London Press, 1973, 196 pp.,
Unibook Edition £1·50, Boards Edition £3·00

Counselling should be a central and integral part of the educational process for all students enabling them to explore their own feelings without fear and to examine their values and objectives without risk of condemnation. This is the underlying theme in a descriptive account of the work of the counsellors in the Appointments and Counselling Service at the University of Keele. The Keele Service embraces educational, vocational and personal counselling and the way in which it operates is examined in detail. Tables are given which provide a valuable insight into the use that is made of the Service, including an interesting analysis of the frequency of certain kinds of presenting problems. In any one year the counsellors reckon to see nearly 50 per cent of the total students in the university. The first part of the book includes an ideal model of the counselling process and a discussion of the role of the counsellor within his institution detailing the ways in which he is likely to relate to the student body and to academic, administrative and welfare staff.

The second part, which is the most valuable and stimulating, concentrates on what actually happens in the counselling process and is illustrated by several accounts of actual interviews. These fascinating insights into how the counsellor works, of what he thinks and feels in a variety of situations, enable one to clearly appreciate what skills he needs and to understand how he is willing and able to accept the responsibility, inherent in the relationship, which commits him to the client's well-being.

Vocational counselling is seen to be not merely concerned with the provision of information or the examination of attitudes and interests but with enabling the person to engage in a process of self-exploration with the object of understanding himself, his values and his aims. As the previous experience of many young people has not required them to take stock of themselves, this self-analysis can generate fear which counselling can help to dispel.

The approach to the treatment of many academic problems such as, lack of concentration, failure to contribute to tutorial groups, poor attendance at lectures, exam anxiety, etc., is orientated towards behaviour therapy, in contrast to the Rogerian approach that the authors generally favour. The importance in this area of co-operation and assistance from academic staff is clearly recognized and examined in detail.

The chapter on personal problems is excellent, particularly for the way in which emphasis is placed on identity crisis difficulties. The desperate need for contact and meaning which many students experience leads some to sex and others to drugs.

The final section, which considers the role of counselling within the changing environment of higher education, is controversial enough to provoke much thought. An appendix surveys the counselling techniques employed at Keele.

This book is an easy to read account of what counselling is all about and can be recommended without any reservation, not only to counsellors, but to all who are really concerned with higher education. DEREK HOPE

The Open University

READING DEVELOPMENT

Improved standards of reading, particularly in the middle school years, are an urgent social need. For everyone directly concerned with reading development, this course highlights means of improving reading competence among children of all ability levels. The course covers reading for learning and enjoyment as well as early and remedial reading. The emphasis is on practical aspects.

Altogether there are fifteen Post-experience courses available at varying prices: (a) The European Economic Community; (b) New Trends in Geography; (c) Methods of Educational Enquiry; (d) Biological Bases of Behaviour; (e) Reading Development; (f) Technology for Teachers; (g) The Handicapped Person in the Community; (h) Environmental Control and Public Health; (i) War and Society; (j) Reformation Studies; (k) Electromagnetics and Electronics; (l) Systems Behaviour; (m) Computing and Computers; (n) Industrial Relations; (o) Systems Management.

The Post-experience Student Office, The Open University
P.O. Box 76, Milton Keynes MK7 6AN

To:–The Post-experience Student Office, The Open University
P.O. Box 76, Milton Keynes MK7 6AN

Please send me Prospectus and Application Forms for 1975 home-based Post-experience courses.

Name (Mr. Mrs., Miss)..

Address..

....

Please insert letter reference
for courses of interest to you ☐ ☐ ☐ PE20

New NFER Books

PIAGETIAN RESEARCH: A Handbook of Recent Studies

Sohan Modgil

This fully comprehensive handbook is sure to prove a major contribution to educational psychology.
Hardback £6·50

Innovation in Higher Education

Edited by K. G. Collier

Six innovatory case-studies in higher education are examined within an underlying theoretical framework.
£2·15

Race, School and Community

Francine Taylor

A strong argument for the allocation of greater resources to the schools, if we are to solve the educational problems of children from ethnic and racial groups.
June, £2·50

Legibility in Children's Books

Lynne Watts and John Nisbet

A fully illustrated review of research on typography as it affects children's books. Clear and readable.
About £2

Environments for Learning

Kevin Marjoribanks

A collection of readings examining a new approach to research in the learning environment of individuals.
July, About £3

Curriculum Development: A Comparative Study

Philip Taylor and Mauritz Johnson

An interesting collection of essays examining the extent and nature of curriculum development in several countries and the indications for the future.
August, About £2·50

NFER Publishing Company

2 Jennings Buildings

Thames Avenue

Windsor Berks SL4 1QS

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £2·75 post free. Overseas (excluding USA and Canada*): £3·20 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: 95p post free. Overseas*: £1·10 post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA and Canada: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Enquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Bucks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the text, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or

2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© *National Foundation for Educational Research in England and Wales, 1974*

Designed by H. A. Senter

Printed by King, Thorne & Stace, Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

121
5

In this issue

Small group teaching methods in higher education
Social class in educational research
Reading ages and tests

[illegible]

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd.

EDUCATIONAL RESEARCH

Volume 17 Number 1
November 1974

Contents

Bloom's taxonomy and the objectives of education *C. P. Ormell* 3

Team teaching in further education *J. Mansell* 19

Interdisciplinary education: lowering the barriers to effective learning
J. P. Wilson and *V. D. MacMurray* 27

Limitations of personality testing as applied to college of education students *F. C. Lamb* 34

Self-concepts among secondary school pupils *B. Thompson* 41

The swing from science: a perspective from hindsight *D. Duckworth* and *N. J. Entwistle* 48

Criteria for the assessment of pupils' compositions *P. Berse* 54

A review of three journals for 1968–70 *T. Derrick* 62

Interpreting research: sampling' – A reply *G. A. Cortis* 66

Book reviews 71

Diary No. 942(3)
Date 25/9/75
d. N
Bureau Edul Psy. Research.

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 17 Number 1
November 1974

Bloom's Taxonomy and the Objectives of Education

C. P. Ormell, School of Education, Reading University

Summary
Bloom's taxonomy, which has been influential in underpinning many of the curriculum developments of the last fifteen years, may be criticized on various grounds. It is a mistake to suppose that Bloom's taxonomy, or any other proposed classification of objectives, can ever be wholly independent of questions of value. On the contrary, it appears that Bloom's taxonomy 'suits' the expression of a pragmatic/materialist ethic, and does not particularly suit the expression of ethics either traditionally, or currently, associated with the concept of 'education'.

In implementation Bloom's taxonomy—it is argued—is a disappointingly blunt instrument. This is particularly apparent when we try to use it in the analysis of mathematics in the upper school. In history the taxonomy seems to lead to conclusions which invert both commonsense and the judgement of experienced teachers. The major criticism of Bloom's taxonomy which is advanced in this paper* is, however, that the taxonomy omits a vital ingredient in education: the development of imaginative understanding. The crucial problem is therefore to identify the behavioural expression of imaginative understanding. This is defined as a 'continuity of fluent response to "if . . . then" . . . questioning'.

In the final part of the paper a new approach to the objectives of education is sketched, based centrally on the concept of imaginative understanding. The problem of formulating a satisfactory account of the objectives of education is, it is noted, particularly acute in relation to the education of the hyper-aware student. The valuable mental residue of imaginative understanding is confidence: and confident response in turn often acts as a sign that the student really understands.

1. The objectives of education

(a) *The implications of a method of classification*

The question, 'What are the objectives of education?' is undoubtedly a *philosophical* question[1]. It is also clearly bound up with value judgements of a very general kind. Nazis, Thomists, Marxists, Bhuddists, Maoists, Zionists, among others, will tend to answer it differently. It seems probable, therefore, that objectives of education, like probabilities, are essentially *relative* to other propositions: in this case to a system of values. For example, the recognized objectives of a system of state education will tend to differ from those of private regimes of education, because the system of values which legitimately can be taken for granted by the state will have less overall definition than that of any particular individual or closely-knit group.

In this situation it is clearly hazardous to adopt a method of classification of the objectives of education without first analysing the values implicit in the classification, and checking that they are not contrary, or even contradictory, to those which the system does, or should, embody. It would be naive to suppose that a 'mere' classification of objectives will necessarily be independent of questions of value: particular classifications, like particular institutions, tend to be suited to the expression of certain values and unsuited to the expression of others.

The mechanism by which systems of classification impinge on questions of value is, in essence, that classifications tend to throw emphasis onto certain qualities, and tend to diminish the apparent significance of others.[2] Hegel's classification of movements of social thought into *theses* and *anti-theses* may have looked innocuous enough when it was first formulated. Surely, it must have been said, it is a tautology that one can always divide movements of social thought into these two categories? Does this not have the certainty of other analytic truths, such as 'It is either raining or not raining'? Yet we now see that the dichotomy subtly prejudges the issue whether, at any given time, there *is* a definite social 'thesis': and it can hardly be disputed that its effect is to encourage a polarization of society into prematurely identified positions.

---

*This is a revised version of a paper which was read by the author at the first of a series of Joint NFER-Reading University, School of Education Seminars on October 30th, 1973.

Similar mechanisms operate when we adopt classifications of many kinds; for example, it is now clearly recognized that the classification embodied in Newtonian mechanics (which makes clear distinctions between space and time, energy and mass, momentum and position) was not capable of being developed beyond a certain point. It had to be substantially modified before physicists could formulate dynamical models suited to phenomena in the 'very large' and in the 'very small'.

It hardly needs to be said, therefore, that it is important to look at a given classification of educational objectives critically: to try to identify the hidden assumptions embodied in it, and to delineate the kinds of value-systems in education to which it may be usefully related. What one cannot do with intellectual credibility is to assume that the business of classifying educational objectives is a purely technical matter, which is independent of the central philosophical questions relating to the aims and values of education.

(b) *Synthesis versus evaluation*

Turning to Bloom's taxonomy of the cognitive domain, we find a way of sub-dividing objectives which puts 'synthesis' and 'evaluation' squarely at the top of the hierarchy. It may be argued that these are self-evidently 'higher' objectives than those of analysis and application.[3] But if we see this order of objectives as 'self-evident', may we not be already presupposing the kind of value-system which fits the taxonomy most neatly? Consider the field of literary performance. Here no one seriously disputes that the evaluator (critic) is of considerably lower stature than the creator (writer). No one would seriously contend that the performance-linked mental operations of the critic embody and transcend those of the writer. The critic may do his work *after* the writer has finished, but in no sense are the objectives of the critic more 'advanced' or more 'significant' than those of the writer. It is a commonplace of ordinary conversation that 'It is easy to be critical': that 'it is much harder to be constructive'.[4]

The same kind of relationship between creativity and criticism may be seen to operate in art, in curriculum development and in science. In research mathematics there is no recognized activity of 'evaluation', except that performed by research boards and the referees of publishers and learned journals. In all these spheres the notion that 'evaluation' is the final educational objective simply does not correspond to the situation. In research mathematics *analysis* (in the Bloomian and ordinary sense) is just as important as synthesis.

'Evaluation' is open to dispute as the final educational objective if only because the value-judgements implicit in outstanding pieces of literary work often 'come over' more powerfully and more impressively than the explicit value-judgements of the evaluator/critic. It is widely agreed that the most effective evaluations of society may be found in the novels of writers such as Victor Hugo, Charles Dickens, George Eliot, Leo Tolstoy, Aldous Huxley, George Orwell. And it seems natural to describe the syntheses achieved by such great writers as 'embodying' profound evaluations of society.[5]

The word 'synthesis' implies more than the mere assembly of disparate bits to form a whole: it implies that the bits fit together with a kind of 'authority': that they form an 'organic unity'. And clearly this cannot be achieved without an acute awareness of what *will fit* well together, i.e. of prior imaginative 'evaluation'. This is the sense in which evaluation is bound up with, and is presupposed by, anything deserving the title 'synthesis'.[6]

The main reason, however, why it is hard to think of 'evaluation' today as the final educational objective is that in any pluralistic society such as ours there is no one pre-eminent system of values within which 'evaluations' can be conducted. And evaluations conducted in different systems of values will often be, in Kuhn's phrase (1965), 'incommensurable'. The Maoist student who evaluates a piece of close economic reasoning as 'tosh', is not doing something comparable to a serious-minded business school student who takes it to pieces in minute detail and comes to the conclusion that it provides the best projection available.

(c) *Behaviourism*

At this point it may be asked whether the fundamental error in Bloom's and other derivative classifications is the behaviourist assumption on which they are based. Is it not absurd to try to classify educational objectives in terms of *behaviours* when the primary objective of education (on most interpretations of education) is the *mind* of the child? This is a complex question on which much that has been said misses the point because it presupposes outmoded conceptions of the nature of psychology. Clearly there is little point in defining educational objectives in terms of 'states of mind' conceived as being intrinsically private states. If it is useful to state publicly that certain objectives are the principal objectives of education, they must *of course* be objectives the achievement of which can be publicly verified. In this sense the behaviourist assumption is unavoidable: if we reject it, we *ipso facto* reject the exercise of trying to state publicly what the principal objectives of education are.

But in 1956 behaviourism was seen less as a necessary presupposition to any scientific approach to psychology, and more as a reaction against the uncritical dualism of the past.[7] It is convenient to call this stage in the development of psychology 'crude behaviourism'. We can see now, with the advantages of hindsight, that crude behaviourism imposed categories onto psychology which did a considerable degree of violence to the subject. An example of this was the rejection by crude behaviourists of the very existence of mental images. Since 1956, research into the nuances of human behaviour has led to the recognition of the behavioural expression of many states previously thought to be wholly 'mental' and therefore private.[8] This has been supported on the one side by a mass of work in linguistic philosophy, which leads us to identify much more explicitly (i.e. behaviouristically) than hitherto what is meant by saying things like 'He *promises* to join', 'She *realized* what would happen', 'He is *mindful* of his responsibilities'. On the other side, it has been supported by impressive work from the neuro-physiologists in identifying, for example, the precise nerve cells which are sensitive to particular bits of an animal's visual field, the effect of hallucinogenic drugs, and the study of the process of dreaming.

The net effect of all this development might be summed up by saying that behaviourism need no longer be crude. Or, to put the same point more usefully, there is now little point is describing a particular brand of psychology as 'behaviouristic' since all psychology will of course consist in the study and explanation of human behaviour.

There is, however, a tendency which is quite widespread today to treat things which can be measured simply, directly and reliably as being more 'real' than things which are difficult to measure directly and reliably.[9] A classic example of this is the scorn with which Bloom deals with the idea of *understanding* as a prime objective of education.[10] According to Bloom this is not a precise idea, and hence not a fit candidate for inclusion as a major heading in the taxonomy.[11] But whether it is a precise idea or not, the word 'understanding' possesses a meaning in ordinary language, and hence is capable of being analysed linguistically by Austinian methods.[12] As a result of such analysis we might reasonably expect to develop a workable (i.e. behavioural) definition of it. After all teachers can very often tell whether or not their pupils understand a particular argument or idea; so they must be aware in some sense of behavioural signs of such understanding.

The important point, however, is that any difficulty which may be encountered in producing a behavioural analysis of 'understanding' is not relevant to the central issue. The central issue is: What are the general objectives of education? To exclude 'understanding' from the very beginning as a general objective, on grounds of imprecision, is to evade the central issue. The fact that we may find it hard to measure 'understanding' by administering tests does not answer the question whether understanding is, or is not, a prime objective of education. One might hazard the guess that a majority of educators throughout history, and particularly today, *have* thought that understanding was a prime objective of education. Almost everyone says that they try to teach 'for understanding'. It is, to put it mildly, extremely odd to brush this issue aside on grounds of technical difficulty. (The point being that the Bloomian exercise ought not to have been attempted at all if the behaviouristic tools needed to implement it properly were not sufficiently well developed to tackle the central problem.) Bloom himself remarks that a 'taxonomy' ought to reflect phenomena as they are, rather than attempt to impose a system of classification onto them.[13] In this central respect Bloom hardly manages to meet his own criteria.

In some passages Bloom speaks as if it were intended that the taxonomy as a whole would provide an analysis of the objective of 'understanding' into its constituent elements. This is an altogether different stance: but to discuss it adequately it is necessary first to have completed a critical survey of particular aspects of the taxonomy. We return to this in Part 2, Section (d).

(d) *Values implicit in Bloom's taxonomy*

We have seen that Bloom's taxonomy allows little room specifically for the classic objective of 'understanding'; but this is, after all, a negative point. What can one say about the approach to education on which the taxonomy exerts the least pressure and thereby allows the greatest room? What, in fact are the values embodied in the Bloom taxonomy?

The answer seems to be that the taxonomy reflects an attitude to education which may be described as 'materialist'. It speaks over and over again of education as the handling of communications. The picture which seems to be permanently present in the background is of the student as handling *material*: acquiring it in bits (knowledge of specifics), turning it over (comprehension), allocating it (application), dissecting it (analysis), putting it together in new ways (synthesis) and comparing the results with other results or established criteria (evaluation).

One's doubts about this perhaps reduce to the feeling that it pre-supposes that no one is asking what education is, after all, *for*. The whole process appears to be instrumental, and to consist in a preparation for a way of life in which the six major categories of the classification will correspond to the main work activities of the educated person. Or to put it in another way, it appears that a student could go through all the processes which correspond to the Bloom headings without achieving any genuine 'understanding'.[14] One gets little sense of the appeal of education for education's sake, and a lot of the feeling of the training of children to be conspicuously effective in exercising skills on tasks of various levels of complexity.[15]

One may thus see the taxonomy as reflecting in a very subtle way the ethos of a product-orientated society.[16] The levels of activity it defines correspond surprisingly closely to the main functionally defined levels in the society:

Type of person	*Type of job*
politician	evaluation
civil servant, director	synthesis
scientist, technologist	analysis
manager	application
skilled man	comprehension
unskilled man	knowledge of specifics

One might describe the taxonomy, therefore, as being predisposed to fit a pragmatic, materialistic, meritocratic view of society, in which few profound questions of value arise, and in which explicit evaluation is always expected and preferred. These look uncannily like the assumptions of the old, pre-crisis consensus in America. But they are hardly those implicit in either the older European concept of a 'liberal education', or in more recent radical theorising about the social aims of 'education for life'.

2. The level of specific application of Bloom's taxonomy

(a) *The problem of implementation*

The analysis of Part 1 was concerned mainly with the outline categories of Bloom's taxonomy: and various reasons were advanced for resisting the implied over-view of educational objectives which it offers. It may be argued, however, that although various criticisms can be levelled at the main headings of the classification, and at the precise order of importance implied by the ordering of the hierarchy, nevertheless the taxonomy undoubtedly *works*.[17] It is certainly true that the examples offered by Bloom in illustrating the taxonomy cover a very wide field, and that the classification is apparently capacious enough to cover anything one might wish to classify.[18] Like the Bible it seems to contain references to everything, and like the Bible one's initial criticism is not directed at lack of comprehensiveness, but at the lack of overall logical coherence. It appears that many of the categories overlap, that the place of a particular behaviour in the hierarchical order will sometimes depend crucially on the prior history of the student, and that in places inversions of the normal hierarchical order may be expected to occur.[19] All this is, to some degree, admitted in the introductory sections of the book[20]: certainly it suggests that a degree of muddle has crept into the taxonomy and that fresh thinking of some kind is needed to sort it out.

The best way to tackle the issues arising here is, I think, to take a close look at the kind of problems which arise when one tries to apply the taxonomy rigorously in one or two subjects. The subject I shall consider first, and in most detail, is mathematics. Afterwards I shall take a brief look at some problems involved in applying the taxonomy rigorously in history.

The key to the discussion is of course the degree to which one expects a *rigorous* implementation. It is not difficult to implement the taxonomy impressionistically in mathematics (as in other subjects), and if one has colleagues with a disposition to form similar impressions, one may achieve a locally agreed view of the situation. But what is immensely difficult is to implement the taxonomy in a way which is not wide open to critical dispute.

The first point which arises is that most problems in pure mathematics at, say, sixth-form level, will consist mainly of *analysis* followed by *synthesis*. The student will be expected to produce a coherent argument leading to his conclusions, or even a formal proof. And to do this he or she will have first to analyse in what the problem consists.[21] The difficulty arises, however, when we try to say what the components of such 'analyses' and 'syntheses' are. The problem is that in mathematics the boundaries between knowledge of specifics, comprehension and application are exceedingly blurred.[22] Very few problems in mathematics in the upper school can be uniquely analysed into these elements. What happens is rather that the student brings a personally unique compound of skills and awareness to bear on the problem. Knowledge of a powerful theorem may dramatically simplify a problem, but the student still needs to realize that the theorem applies legitimately in this particular instance. Thus knowledge and application are intrisically bound together.[23] To know a theorem in mathematics and not to be able to apply it, is not to 'know' it in any significant sense. Moreover there are a great many 'adjustment' steps in mathematics which need to be handled with skill. These steps can be regarded as examples (i.e. *applications*) of general laws, principles and axioms: or they can be regarded as examples of *comprehension*.[24]

The overall conclusion to which one is led here is that the Bloomian categories of behaviour are extremely amorphous in relation to mathematics. They cut across the natural grain of the subject, and to try to implement them—at least at the level of the upper school—is a continuous exercise in arbitrary choice. The result is that at the end of such an analysis one does not feel that the natural behaviourial structure of mathematical performance has been revealed, but that an awkward classification has been imposed upon it.

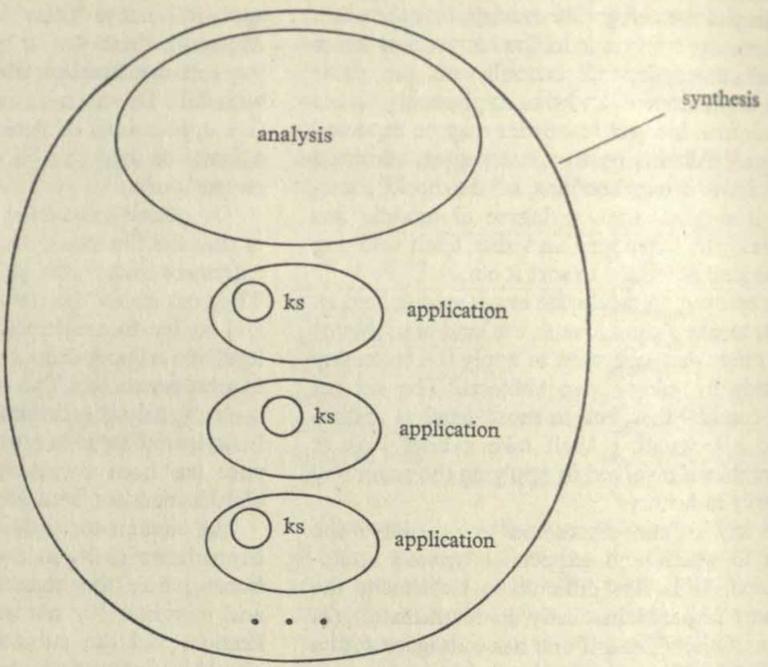
The attempt to apply the taxonomy rigorously in mathematics is, in the author's view, crucial: because here the contrast between the blurred and unsatisfactory nature of the imposed classification, and the subject matter itself, is most clearly seen. One feels that it ought to be possible to analyse the behavioural objectives in a formal subject precisely. Instead one finds a glaring mismatch between the (lack of) precision of the classification and the precision of the subject being classified.

Probably the best complexion which can be put on the situation from the point of view of the taxonomy, is to say that the analysis of almost all mathematics problems at the upper school level has the same kind of structure. The overall behaviour type is one of *synthesis* composed of an *analysis* stage followed by an *application* stage (with some embedded knowledge of specifics). See Figure 1.

One's final complaint is therefore that the taxonomy fails to reflect sufficiently sensitively the range of different behavioural patterns which actually occur. It treats everything alike and thereby achieves a zero operational effect.

In essay writing in history, on the other hand, we find the analysis of behaviours into *syntheses*, composed of *evaluations* of policies, *analyses* of motive, *applications* of social, economic and political principles, *interpretations* of documents and *knowledge* of specific facts. This may sound like a much more satisfactory state of affairs. But . . . is producing evidence of an *under-*

FIGURE 1



standing of the Council of Trent a simpler, more primitive behaviour than *evaluating* Hess's flight to England? Is *interpreting* Constitutional documents of the 15th Century a more primitive, less complex behaviour than *analysing* the motives of Nero? Is a mention of the European inflation which resulted from the influx of Spanish gold in the sixteenth century a knowledge of a specific fact, or an *application*? The conditions for 'application', as defined by Bloom, seem to require that the student recognizes an instance of a principle of this kind without prior prompting. But how is this possible in such a well-trodden field as history, in which interesting instances of applications of general principles are not very numerous, and hence almost invariably become part of the common currency of the subject (i.e. graduate into knowledge of specifics.).

In mathematics, it might appear at first that we have the ideal subject for implementing the taxonomy: a subject which relies little on knowledge of specifics and consists largely in the exercise of intellectual skills. However the taxonomy appears ill-adjusted to the level of discrimination required in mathematics. In history, on the other hand, a subject apparently largely based on knowledge of specifics, there seems to be much more scope for implementing the taxonomy. The trouble here is, however, that the classification of behaviours which results hardly corresponds at all to the realities of historical education. Evaluating Hess's flight to England is a *much* easier activity than trying to assimilate the meaning of the Council of Trent. A taxonomist would have to be clever indeed to convince himself that the principle of organization by complexity is exhibited in these official Bloomian precedences.[25]

(b) *Modifying the taxonomy*

An interesting question at this stage is what modifications one could recommend to be made to Bloom's taxonomy in order to turn it into a more effective classification device. Part of the answer seems to be that we should abandon any pretence that a principle of organization by complexity is involved, and split the taxonomy, in effect, into six parallel taxonomic categories.

Bloom distinguishes different levels of knowledge of specifics, but seems to fail to notice that for each such level there are equivalent levels of comprehension. Nor is application just a simple affair of instances of laws and principles. Here, too, we find a complete hierarchy of rules, laws and principles of different levels of generality and hence of different levels of application. (It is the neglect of these distinctions that renders Bloom's taxonomy such a poor instrument in mathematics.) In terms of analysis the same levels occur. It is one thing to analyse 273 into its prime factors; yet another to analyse the natural modes of vibration of a rigid three-dimensional object. Synthesis, too, clearly comes in different sizes. John who is inventing a $3+3$ magic square is showing a behaviour correctly classifiable as 'synthesis': but his behaviour is obviously of a very different order of complexity from that of Jack who is designing a buffer store for the school's homemade computer terminal.

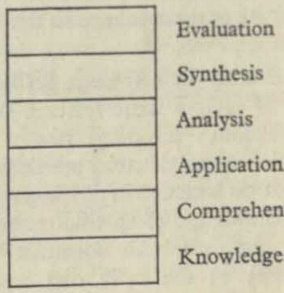
To develop the taxonomy usefully on these lines, however, some modification of the definition of the knowledge category would seem to be needed. It is just as easy to memorize the symbols '$F=\mu R$' as it is to memorize the symbols '$E=mc^2$' but the former expresses a result on the lowest level of scientific law, whilst the latter expresses one on the highest. Clearly a capacity to recall meaningless formulae ought not to be treated as any kind of constituent of education (not even a low-level one).[26] A minimum alteration here would surely be to define the knowledge category in terms of 'awareness of what is the case'. Such awareness would need to be tested in an appropriate way on each level. One would not assume that a boy or girl who could merely recall '$E=mc^2$' possessed anything legitimately describable as an awareness of the meaning of Einsteins' law.[27]

In addition to these distinctions it seems apparent that a 'weight' distinction is needed: one can have a problem, or a curriculum, which contains material requiring a lot of comprehension on level 3 or a little comprehension on level 3. It is not enough to say 'this problem requires comprehension on level 3': children have varying degrees of grasp, confidence and stamina. It is unrealistic to suppose that a child can either operate indefinitely on level 3 or not at all.

Finally we also seem to need a 'visibility' distinction. Instances of general rules are rarely of constant visibility. For example, a^4–1, a^4–b^4, a^2–1, a^2–b^2 are all instances of 'differences of two squares' and can all be factorized by the same rule, *viz.* $X^2-Y^2=(X-Y)\ (X+Y)$. But they are not equally *visible* instances in which this rule can be applied. In mathematics visibility in this sense plays a central role, and any attempt to do an analysis of behavioural objectives which does not take this into account will obviously not get very far.[28]

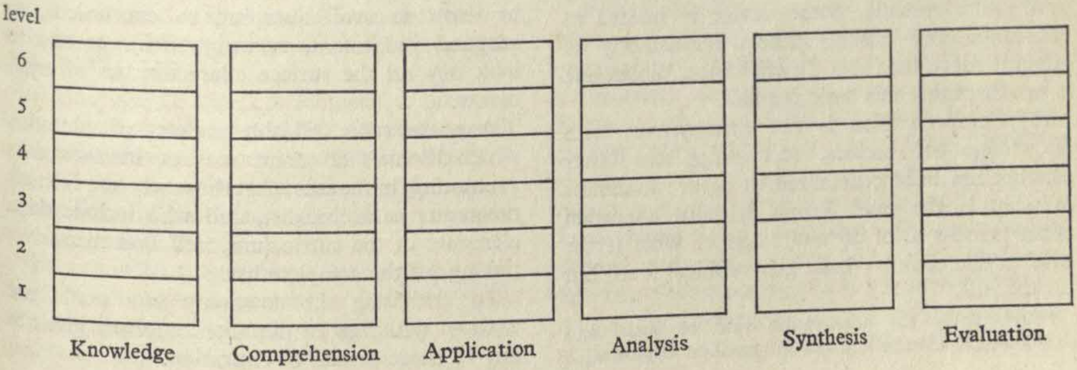
Using these additional instruments (and assuming that the massive subject-classificationary leg work which they presuppose had been done),[29] we could, I think, do many of the jobs with the Bloom classification which one would like a classification to do. Instead of adopting a pattern like this:

FIGURE 2



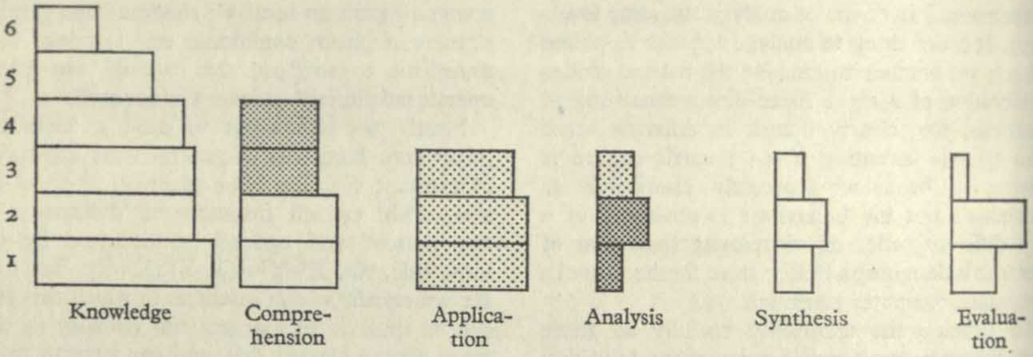
as the Bloom taxonomy proposes, we would adopt one like this:

FIGURE 3a



A 'taxonomic profile' of a particular examination paper might look like this:

FIGURE 3b



Width of the block denotes weight, shading denotes lack of visibility.[30] The question still remains, however, whether the headings of the six blocks and their horizontal subdivisions are the most appropriate ones to use. What *are* the objectives of education? In looking at questions concerning the technical efficiency of Bloom's taxonomy we have temporarily set aside the discussion of fundamental issues begun in Part 1. We now turn back to these in an attempt to sum up the state of the argument so far.

(c) *Evaluating the taxonomy: positive features*

In this section some features of education to which Bloom's taxonomy draws our attention, and which seem valuable, are listed: these ought surely to be included in any improved taxonomy.

(1) Knowledge of specifics tested by recall of verbal and symbolic formulae is hopelessly inadequate as the main diet and objective of education. Recent changes* have made this conception of education completely obsolete in the form in which it was commonly practised until quite recently. Some device is needed to encourage and enable schools consistently to attempt more than this. In criticising Bloom one is not disputing this basic premise.

(2) Comprehension is not synonymous with knowledge of specifics, even when the latter concept has been generalized to cover 'awareness of what is the case'. People without education often possess a lot of 'awareness of what is the case' at the ordinary basic geographical level, but this does not imply that their comprehension of given geographical information is good. Here again one is not disputing the need for such a distinction.

(3) 'Application' covers the understanding and use of generalizations of all kinds throughout the fields of science, technology and affairs. This must clearly be of central importance, especially as subject-based levels of generality will, it seems, need to be built into the structure of any improved taxonomy from the beginning.

(4) 'Analysis' covers the need for critical concentration; for the need to stand back and see the overall shapes of things at a lower level of definition. These are certainly behaviours most of us would wish to encourage.

(5) 'Synthesis' covers the need for students to be able to argue and present a coherent statement of a case. Constructive traits of thinking and some element of creativity are also widely considered important. It seems to be important on any analysis to distinguish powers of expression and synthesis proper.

(6) 'Evaluation' covers the need for students to learn to avoid jumping to conclusions, to suspend judgement, to weigh evidence and to look beyond the surface characteristics of arguments.

These are some valuable aspects of education which Bloom's taxonomy may be interpreted as promoting, in the sense that those who use Bloom's taxonomy as a checklist, and who include these elements in the curriculum, will find themselves ticking off the complete list.

In criticising the taxonomy one would not wish to withdraw or diminish emphasis given to any of these aspects of education.

---

*Particularly the accelerating rate of social and technological change and the information explosion.

(d) *Evaluating the taxonomy*: *negative features*

What, then, is wrong with Bloom's taxonomy: what is the basis of the criticism?

(1) The banality of a value-system in which 'evaluation' is placed above 'synthesis'. If a simple, crude, ordering of objectives were to be attempted it would be natural to place 'synthesis' at the top; 'synthesis' being interpreted in the ordinary sense. This is the order of priorities which, one suspects, most educationalists actually observe. The case for this was argued in Part 1 (a).

(2) Over-emphasis on the cognitive-affective distinction. Here the boundary is much less clear-cut than the taxonomy implies: the comparative irrelevance of the classification of affective objectives in most practical applications of the taxonomy underlines this point.[31]

(3) Over-emphasis on the cognitive-psychomotor distinction. That these are not independent categories is suggested by the development of apparatus work in primary mathematics.[32]

(4) The taxonomy fails to distinguish subject-related levels of the upper five categories, and fails to make weight and visibility distinctions. The case for these distinctions has been argued in Part 2 (b) above.

(5) The epistemological naivety of the process-content distinction on which the taxonomy is based. The point here is that the process-content distinction implies a clear-cut boundary between that which one cannot influence (content) and that which one can (process). This overlooks the fact that behaviours falling under the higher objectives must—if they are to make sense—be based on *true* mappings (comprehension), *true* laws (application), *true* dissections (analysis), *etc.* In other words the principles and rules of 'process' which we hope our students will learn to use, must also reflect 'what is the case'; and for the student to know that it *is* the case is for the student to know something specific about the invariants of the real world.[33]

(6) Omission of the development of imagination as a major constituent of understanding and a major objective of education. This is connected with the rejection of mental images in the crude behaviourism of the fifties. We shall discuss this in more detail in Part 3.

We saw in Part 1, Section (a) that the taxonomy, taken as a whole is sometimes interpreted as constituting an analysis of the allegedly vague, ill-defined objective of 'understanding'. Criticisms (4), (5) and (6) above and the counter-example of Note 14, show that this claim is less than plausible.

Taking the full range of its deficiencies into account, the notion that patching up the taxonomy will produce an instrument which will work effectively, does not appear at all satisfactory. A much more radical reconstruction seems to be needed; which, (i) will relate to conceptual levels in subject areas, (ii) will not imply a single linear hierarchy of objectives, and (iii) will include the ingredient of imaginative redevelopment from the beginning.

In Part 3 we turn to look at the outlines of a new approach to classifying the objectives of education embodying these elements.

3. The objectives: a new approach

(a) *The subject matter of education*

The problem of the objectives of education is clearly a difficult, elusive and massive one. A useful first step in addressing it is to divide the problem into two. First there is the problem of objectives of education for the youngster who is aware to an extent *greater* than average of the world around him, its problems, concerns and meanings. Second there is the problem of objectives of education for the youngster who is aware of the world around him to an extent *less* than average. The two problems are initially somewhat symmetrical. The first youngster needs to make sense of the world around him: he needs what one may call a 'Forsterian' kind of education. The second youngster needs to be shown the nonsensicality of the mental blinkers in which he is enclosed: he needs to be led to be more aware of the size and depth and interest of the world.

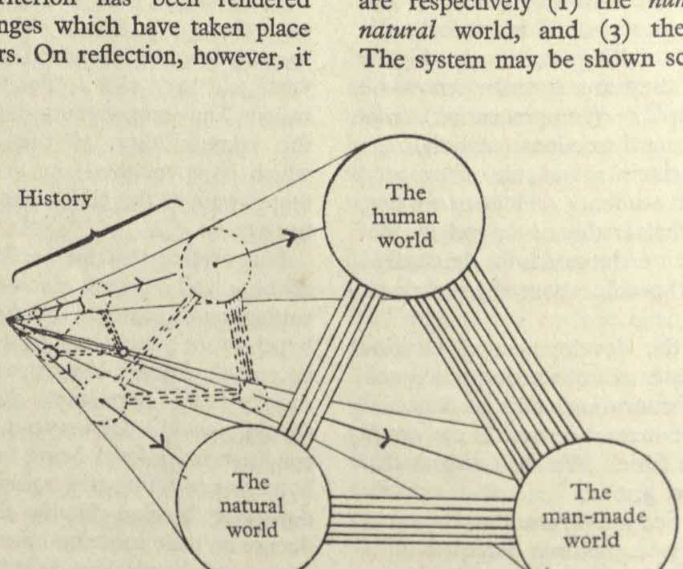
It is evident that for social reasons youngsters of the second kind are much more numerous than youngsters of the first. Nevertheless, because the hyper-aware youngsters of a generation tend to become the leaders of that generation, it is clearly socially very important that their education should 'work'. Observation shows that this condition is far from being satisfied today. Many hyper-aware youngsters today suffer from breakdowns or become totally alienated. Their influence on their contemporaries is such that many others, not themselves initially in this category, acquire the same attitudes. If we introduce a weighting factor to reflect this influence,[34] we end up with two groups of roughly similar total weight: thus it is reasonable to treat the two sides of the problem of objectives with the same degree of seriousness. One of the underlying problems today, the author suspects, is that in correcting the imbalance of the past in educational priorities, we have swung over to a position in which serious thought is mainly being given to the second group.

Indeed many recent theorists write as if education largely consisted in stimulating children to learn things, and to become involved and interested in activities outside the original ambit of their backgrounds. This may express quite well a general objective suited to the needs of the second group of youngsters, but it is surely hopelessly inadequate as a prescription for the first.[35]

At this point we must ask ourselves what are the characteristics of the 'educated' man or woman. The answer will not be someone who knows a variety of unconnected things or who has learnt them unsystematically or wholly intuitively. The essential characteristic is that he or she should possess a *connected* view of the world: an awareness of things in proper perspective: an absence of irrational mental blocks. This is summed up in E. M. Forster's famous phrase 'only connect'.[36] It could be argued of course that this criterion has been rendered obsolete by the changes which have taken place in the last sixty years. On reflection, however, it appears that the general objective of connectivity is more relevant today than it was in 1910, the year in which *Howard's End* was published. Just because explosive innovation has bafflingly altered many relationships (social, technological, scientific) between disciplines, between facts, and between ideas, it is now all the more difficult to 'connect'. And it is all the more important that education should lay the basis of a dynamic, but connected, view of the human condition.

One aspect of this is that the view of the modern world which we present in, or which emerges indirectly from, the curriculum, should be seen to be firmly connected with the world of yesterday and of earlier times. This is a further ground for doubting the wisdom of a wholly 'process'-orientated approach to the objectives of education. Another aspect of this is that to establish a prior basis for connectivities we need an adequate provision of early emphasis on different areas of human experience: i.e. in the phrase of Working Paper 45 'a balanced curriculum'.[37] What are the main areas which need to be as it were, 'laid down' early on, so that innumerable connections can be made later without distorting the underlying framework of ideas which will have evolved in the youngster's mind? The answer seems to be that there are three main areas: three main 'worlds', a degree of understanding of which we need to cultivate. These are respectively (1) the *human* world, (2) the *natural* world, and (3) the *man-made* world. The system may be shown schematically thus:

FIGURE 5



Clearly the three 'worlds' are inter-connected, and each in turn is connected with earlier, historical, states of the three-cornered system. Space does not permit a detailed recapitulation of even the main ways in which the three 'worlds' interact, or a lengthy justification of this structure as an outline pattern of the space we need to make available for the connectivities which should be progressively developed in a balanced education. The key issue is, surely, that the student should acquire a degree of understanding of the main elements operating in the three 'worlds'. This seems to be a useful way of clarifying a view of the objectives of education which is, one suspects, coming to be quite widely held today.[38]

It differs from classic statements of the nature of a liberal education[39] mainly in taking the man-made 'world' seriously. It is a scandal, in the author's view, that the man-made world, which is so largely a determinant of fundamental changes in human history, has had so little expression in the usual academic curriculum. We have, in the past, grossly neglected subjects such as industrial archaeology, the history of invention, the history of weapons, technology of all kinds, social implications of technology, applicable mathematics,[40] applications of science and even, indeed, engineering. With so many points missing from the average youngster's knowledge map, it is hardly surprising that many youngsters emerge from the system with a sadly undeveloped and stunted sense of the connectivities of the world.

One of the strengths of Bloom's taxonomy is that it brings some of the characteristic value-judgements of the man-made world to bear on education. But how applicable are the Bloomian categories to the natural world and the human world? Surely in both these worlds we need to place much more emphasis on the passive virtues of understanding, receptivity, sensitivity. Science, as the attempt to understand the natural world, would hardly make any progress if it did not emphasize patience, open-mindedness and a willingness to consider the full range of known facts. There is of course a major place for analysis, but only the very few will get as far as contributing to the synthesis of scientific theory. In the human world, of course, evaluation can be operated from the lowest level; application of moral principles is possible over a wide sector (but is impeded by current moral uncertainties) and synthesis hardly seems to arise. Yet the main burden of education in relation to the human world ought surely not to be a constant evaluating or moralizing about the past, but rather acquiring a body of knowledge about, and understanding of, it.

What are the conclusions of all this? They are, surely, that much more attention needs to be given in our talk about objectives to the central objective of *understanding*; and that the content side of education is under-represented in Bloom's taxonomy. On the other hand content is not the only component of understanding, and a purely passive form of understanding would, in any case, be hardly suitable for anyone except a student who sought a career as a Buddhist monk!

In the early stages of education the elements of what may be called 'soft' competence (i.e. basic types of intellectual competence needed by everyone), psychomotor skills, 'soft' creativity[41] and imagination[42] need to be combined with a lot of exposure to facts about the world to generate understanding. In the later stages, understanding becomes the starting point for hard competences of all kinds (i.e. work aimed at various professions) and for hard creativity in science and technology.

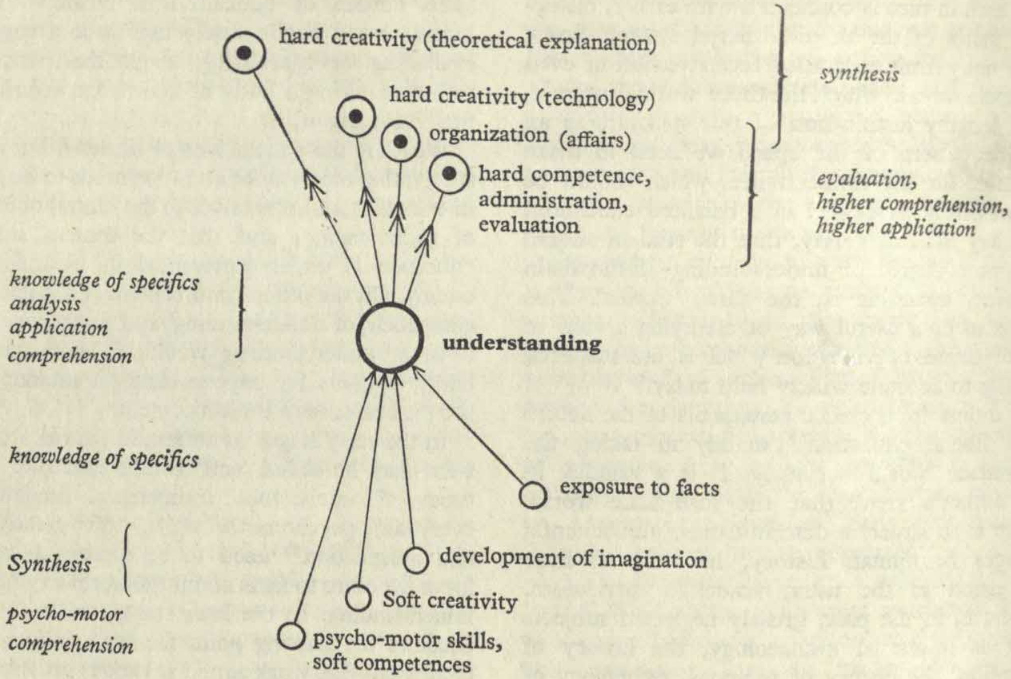
The objectives suggested by this approach may be shown schematically, see Figure 6.

The central problem to be faced now is what is meant by 'understanding': what is the behavioural sign of this? It is sometimes argued—to revert to an earlier point—that Bloom's taxonomy provides an accurate breakdown of 'understanding' into its constituent elements, and therefore that the views presented in this paper need differ little in essence from those of the authors of Bloom's taxonomy.

(b) *The problem of understanding: the importance of imagination*

That there is a close affinity between the objective of connectivity and the objective of understanding hardly needs to be laboured. At a rough level of description one might say that understanding X consists in knowing how X relates to, or connects with, other aspects of the world.[43] 'Knowing' will in many cases, however, be too strong a word to express the psychological situation. Hence the attraction of a weaker definition in terms of 'awareness' of the connectivities of X. Someone who is aware of the main connectivities of X will be someone who

FIGURE 6


hard creativity (theoretical explanation)
hard creativity (technology)
organization (affairs)
hard competence, administration, evaluation
synthesis
evaluation, higher comprehension, higher application
knowledge of specifics
analysis
application
comprehension
understanding
knowledge of specifics
exposure to facts
Synthesis
development of imagination
Soft creativity
psycho-motor
comprehension
psycho-motor skills, soft competences

Note: some roughly corresponding Bloom headings are shown in italic.[44]

could be correctly said to 'understand X'. The question still arises, however, what will count as the 'main' connectivities. The quick answer is that each X will have its own characteristic 'main' connectivities. Let us consider a specific case: one who could be correctly said to 'understand the differential of a car' need not know its history, the metallurgy of the gear wheels, the main sources of supply of parts, etc. But he *does* need to know the number, nature and structure of the main parts, and the kind of dynamic, functional relationship which exists between them when the car is rounding a bend in the road. There are certain thresholds here:

(1) one who did not understand a differential would not perhaps even know what to expect if he opened-up a back axle of a car,

(2) on the other hand he might have learnt the composition of the device by heart and yet have no idea how it fitted together, or

(3) how the parts moved relative to each other when the car was rounding a bend.

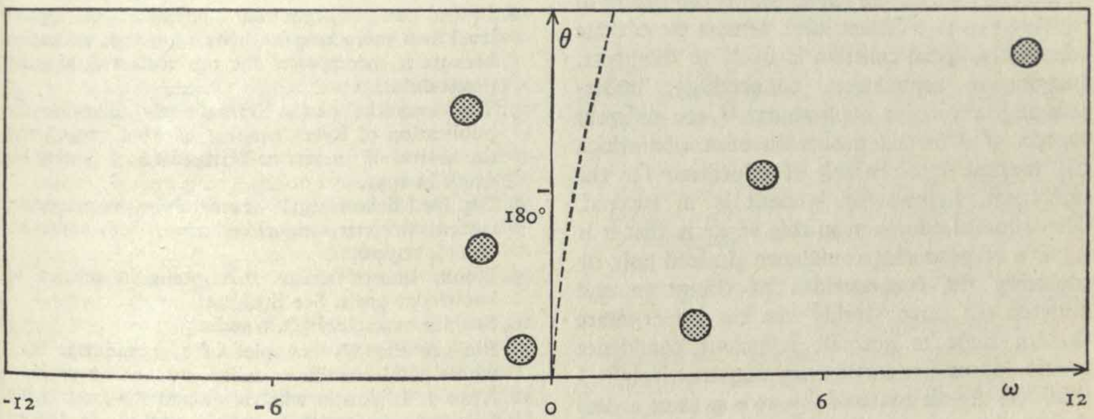
The most unambiguous way to test for understanding of the device would be to put out a series of, say, twenty trays containing plastic gear wheels and other simple components (all fairly similar) including parts corresponding to, but not identical with, the four main components of a differential. One would then ask the student to choose four of the pieces and to use them to assemble a working equivalent of the device. It is not, however, normally necessary to go as far as this in order to test for understanding. The essential element in such tests is whether the student knows the answer to a range of questions of the type 'if X, what then?' In the above example 'X' would cover the cases: (a) the casing is taken off the device, (b) the main components are provided for assembly into the device, (c) the device is operated in a car with perspex covers so that one can see the pattern of relative movement.

In other words the test for understanding is whether the student can answer a limited, but continuous, connected range of 'if . . . then . . .' questions about the situation.

The 'continuity' of the range of questions is an essential part of the criterion. A student might even know by heart what was the positioning of the main parts in six special configurations, and yet fail to *understand* the device. On the other hand a student who understood the device would be able to imagine the relative motion of the parts, and so would be able in principle to answer an indefinite number of questions about what happens.[45]

The situation may be shown schematically:

FIGURE 7



ω = angular velocity of the pinions in revolutions/sec

θ = angle the crown wheel makes with the forward direction of the car

A typical question: what does the dotted curve represent? The student who 'understands' will be able to draw or describe the situation for each value of θ and ω in the range shown, whilst the student who merely knows 'what is the case' in each of the six marked states, will have a knowledge of specifics which only amounts to a facility to answer a tiny fraction of possible questions of this kind.

Clearly the difference between the two students is not that the student who understands the device has absorbed a much greater quantity of information than the other. What he has done is rather to have *imaginatively* assimilated the information he has been given: so that his imagination is able to re-create the state of the device at any point in the rectangle shown.

'Understanding' therefore involves a compound of awareness-of-what-is-the-case and imagination. The behavioural test is a *continuity* of facility to answer appropriate 'if . . . then . . .' questions. Such a behavioural analysis of understanding applies equally to subjects of all kinds right across the curriculum. A student who has imaginatively assimilated *Lear* has achieved something educationally valuable which is merely caricatured, not approached, by a second student who can correctly answer numerous questions of fact about the play, but has failed to understand it.

What are the practical implications of all this? They are surely that we should drop the 'digital' approach towards content in education, and adopt instead a 'holographic' principle.[46] We should not assume or act as if students significantly know anything about a subject until they have achieved a sufficient level of connectivity to make imaginative assimilation possible. The method of teaching needed is one which embodies imaginative energy: we need continually to probe, explore, re-create, using the basic technique of interesting, relevant 'if . . . then . . .'questioning. R. G. Collingwood defined history as the imaginative re-creation of the past,[47] today anything less than this is less than genuine education.

(c) *The value of education*

It has been remarked that prior to the age of popular education the typical educated person belonged to a definite social sub-class: the 'clerical' class.[48] Education provided the passwords of the class. Social coherence sprang from religion, shared assumptions, a common curriculum and a deeply conservative social structure.

Today only remnants of the feeling and style of the old 'clerical class' remain. Most people can now see plainly enough that the scholasticism of the old conception of education is inappropriate and indefensible. Imaginative assimilation, on the other hand, was an ideal formulated by men, such as Collingwood, who tried to identify the highest characteristics of the education of their day. Today we can no longer afford the luxury of treating this as a distant ideal. Unless we educate coherently, social cohesion is likely to disappear. Imaginative assimilation, connectivity, understanding, projective applicability,[49] are different aspects of a certain quality in education which has become indispensable if education for the intelligent, hyper-aware student is to succeed. The value of education in this sense is that it is capable of producing confidence. Indeed only by exploring the connectivities of things in and between the three 'worlds' can the hyper-aware student hope to generate sufficient confidence to use his or her intelligence constructively.[50] I think we are all aware of this to a greater or less degree: because we often accept the youngster's confident 'of course' as a sign that he or she really understands.

Notes

[For convenience B. S. Bloom is referred to throughout as if he were the sole author of the taxonomy. The two volumes are denoted B(1) and B(2); see Bloom (1956, 1964).]

1. Because to tackle it we need to bring to bear on it a kind of systematic, discursive reflection. What is the underlying significance (in terms of aims), of the kind of policies, curricula, methods, courses, *etc.* we commonly judge to be 'educational'? The 'we' here is not one with which *everyone* will be able to identify: 'educational' is an essentially contested concept in the sense which A. P. Ryan (1974) (following W. B. Gallie) has defined.
2. Behaviours which do not fit conveniently into the classification are broken up, and thereby apparently diminished in weight.
3. Although Bloom consistently speaks of the classification as a 'neutral' device (B(1), 14-15), based on the principle of complexity (B(2), 10-11), he also says that 'the order and arrangement of the classes and subclasses expressed a fundamental relation among the possible outcomes of education' (B(2), 10).
4. Bloom's own first illustration of 'evaluation', under heading 6.20, (B(1), 200) shows how elementary an evaluation behaviour may be. There is no sign here of the complexity which is supposed to be the principle of organization of the taxonomy.
5. The point here is that achievements such as those of the writers mentioned have always been regarded as amongst the highest achievements of the educated mind. Compared with these achievements the performances of the literary critics who reviewed and 'evaluated' their work are comparatively commonplace. Administrators, sociologists and businessmen may naturally use a scale of values in which the final objective is evaluation, but this is not the value-system usually associated with education.
6. Indeed one can argue that synthesis at the highest level is a more *complex* behaviour than evaluation because it encompasses the pre-evaluation of many possibilities.
7. This was the period immediately following the publication of Ryle's *Concept of Mind* (1949), and the revival of interest in Wittgenstein following his death in 1951.
8. E.g. Dr Michael Argyle's research on eye movement patterns in conversational behaviour. See ARGYLE, M. (1969), 105-10.
9. Bloom himself makes this point in relation to knowledge goals. See B(1), 34.
10. See, for example, B(1), 1 and 15.
11. But see Bloom's examples of 1.31 and 1.32 B(1), where 'understanding' unaccountably creeps in.
12. After J. L. Austin who developed the most highly focussed version of linguistic analysis in Oxford in the early Fifties. Austin's method may be seen in operation in his posthumous *Sense and Sensibilia* (1962).
13. See B(1), 17.
14. One can see this clearly if one postulates an extinct, un-literary, ancient language used by a society of no intrinsic interest (e.g. one of the barbarian hordes encountered by the Romans). One could imagine the study of such a language being promoted in schools, and youngsters being urged to learn its words (knowledge of specifics), comprehend its sentences, apply its syntactic rules, analyse passages, write compositions in it and evaluate the results. In other words all this conspicuous educational activity might take place, yet be based on worthless material. Clearly there would be nothing *educational* about such an exercise.
15. See J. H. Plumb (1964), 24-44. Plumb complains that when history is treated as material on which to train the mind very little sense of the significance of history in its own right remains.
16. Compare IBM's famous advertising slogan of the sixties: '*ANALYSE: SIMULATE: EVALUATE: CONTROL*'. If one replaces 'simulate' by 'synthesize a model'—which is what is meant—one has a clear echo of the taxonomy. See IBM (1966).
17. Any classification which is not wholly inept 'works' in the sense that one can use it to divide-up phenomena (c.f. Hegel's thesis-antithesis distinction). The question is whether the pattern of division thus obtained is helpful and fruitful of further insight.

18. There is, however, a serious omission in that the taxonomy omits mention of imaginative development which is shown behaviouristically in a *continuity* of fluent response to if . . . then . . . questioning. The point is developed in Part 3.
19. See B(1), 185.
20. In B(2) 11 the question is even posed whether the taxonomy is a 'true taxonomy'.
21. Surprisingly enough this mode of analysis, analysis-of-problems, is not treated by Bloom as an important sub-variety. See B(1), 144-9.
22. The essential skill which students continually exercise in mathematics is pattern recognition. That *a* is an instance of *B* is an example of *comprehension*, a specific mathematical *fact*, and it can also be regarded as an *application* of an infinite class of general rules.
23. This is an instance of what Sockett calls the 'epistemological naivety' of Bloom's taxonomy. See SOCKETT, H. (1971), 23.
24. The basic problem here is that Bloom invariably uses as his examples *single-step* exercises. In a single step exercise the context can be used to decide which taxonomic category is appropriate. But in mathematics each step generates further steps which *a fortiori* have no determinable context in the original 'communication'.
25. *c.f.* B(2), 10: 'We were of the view that the principle of complexity which had become our basis of arrangement was a reflection of the order of difficulty of the learning of the different objectives'. Gabriel Chanan (1974), draws attention to the difficulty of establishing orders of complexity in the humanities.
26. It is an unfortunate side-effect of Bloom's approach that he adopts a low interpretation of 'knowledge of specifics', and then proceeds to recognize it in the taxonomy. A similar point is made by Richard Pring (1971), 89.
27. This is typical of knowledge of specifics generally. It always presupposes a framework of meaning and in the absence of this framework is virtually useless. See SOCKETT, H. (1971), 23.
28. Literally, of course, the 'visibility' is a characteristic of the material, or content, not of the behaviour. The behaviour consists of the ability to recognize instances of varying degrees of visibility.
29. The 'leg work' consists of classification of conceptual levels in the various subject areas.
30. Notice that the visibility factor only affects the categories of comprehension, application and analysis.
31. The 'over-emphasis' is that implied by the production of separate sub-taxonomies for the two areas. The notion that the two areas are 'independent' in any strict sense hardly stands a moment's examination. It is a premise of much recent curriculum development that there is an inter-action between the two which can be used to stimulate cognitive development.
32. In other words psychomotor development of certain kinds tends to facilitate cognitive development. There is a lot of interaction here, especially in art, technology, science and mathematics.
33. When this is lost sight of the danger arises that students will acquire diminished respect for the less obvious constraints. This can all too easily lead to a 'manipulative attitude' towards material in general. The classic example of an intelligent man who acquired a 'manipulative attitude' in this sense is Albert Speer. The production of people like Speer, it has been said, is characteristic of the disassociated scientific/technological education of the Twentieth Century.
34. The weighting factor can be, in fact, just a count of the total number of youngsters likely to be influenced by individuals of the first group.
35. In what follows I shall be concerned only with the problem of the objectives of education in relation to the first group. This is not to deny that much also needs to be done on the problem of objectives for the second group.
36. Forster prefaced his novel *Howard's End* (1910) with this quotation. Some writers e.g. Peters (1973a) stress the possession of 'love of learning for its own sake' as the essential mark of an educated person. This is of course the process by which people become 'educated'. It seems important, however, to distinguish the characteristics of an 'educated' mind as opposed to an 'uneducated' one, and the process by which these characteristics have been acquired is not strictly relevant to *this* question. For example, one would say that Pascal and Wittgenstein were 'educated' both before *and after* they renounced mathematics. Peters (1973b) partly recognizes the force of this point: See 261 and 264-267.
37. See Schools Council (1972).
38. *E.g.* Sir Eric Ashby has urged that we need a form of education for technologists which will concern itself with the social implications of technology as well as 'pure technology'. See ASHBY, E. (1958), Chapter 4, II.
39. See HIRST, P. H. (1965).
40. See ORMELL, C. P. (1973), 222-3.
41. Two types of creativity may be distinguished: see ELLIOTT, R. K. (1971a, 1971b).
    It is convenient to call the new creativity 'soft creativity' and the older 'hard creativity'.
42. See C. P. Ormell (1971), for an account of this in relation to mathematics. For a more general statement of the value of imagination in education see WARNOCK, M. (1973).
43. Richard Skemp (1971) uses the phrase 'schematic understanding' to express this. Unfortunately in his implementation of the idea he has confined his attention mainly to connections inside the formal terminology of mathematics.
44. Headings higher on the page in this diagram are regarded as 'higher' objectives, though there is no simple linear progression as in Bloom.
    Clearly the scheme shown in the diagram reflects a different value-system from Bloom. This is shown by the changed order of the roughly corresponding Bloomian labels placed alongside.
45. *c.f.* Wittgenstein's famous discussion of understanding in *Philosophical Investigations* (1953), beginning para. 155.

46. In other words information does not have, and should not be presented as having, hard edges: it can be progressively extended and enriched.
47. See Collingwood, R. G. (1939).
48. See Gellner, E. (1964).
49. See Ormell, C. P. (1974) 1 and 2-9 for a definition of 'projective applicability'. It means roughly 'the applicability of mathematics to the discussion of possibilities'.
50. The 'second group' of Part 3, Section (a) of course do not need additional confidence: their main need is for intellectual stimulation. Hence the serious mistake involved in treating the problem of the objectives of education as a single problem.

References

Argyle, M. (1969). *Social Interaction*. London: Tavistock/Methuen.

Ashby, E. (Sir) (1958). *Technology and the Academics*. London: Macmillan.

Austin, J. L. (1962). *Sense and Sensibilia*. London: Oxford University Press.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. and Krathwohl, D. R. (1956). *A Taxonomy of Educational Objectives, I*. London: Longman.

Bloom, B. S., Krathwohl, D. R. and Masia, B. B. (1964). *A Taxonomy of Educational Objectives, II*. London: Longman.

Chanan, G. (1974). 'Objectives in the humanities', *Educ. Res.*, **16**, 3, 198–205.

Collingwood, R. G. (1939). *An Autobiography*. Harmondsworth: Penguin.

Elliott, R. K. (1971a). 'The concept of creativity', *Proceedings of the Philosophy of Education Society of Great Britain*, **V**, 1, 97–104.

Elliott, R. K. (1971b). 'Versions of creativity', *Proceedings of the Philosophy of Education Society of Great Britain*, **V**, 2, 139–52.

Forster, E. M. (1910). *Howards End*. Harmondsworth: Penguin.

Gellner, E. (1964). In Plumb, J. H. (Ed.) *Crisis in the Humanities*. Harmondsworth: Penguin.

Hirst, P. H. (1965). 'Liberal education and the nature of knowledge'. In Archambault, R. D. (Ed.) *Philosophical Analysis and Education*. London: Routledge and Kegan Paul.

I.B.M. (1966). *The Scientific American*, **215**, 3, 105–8.

Kuhn, T. S. (1962). *Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Ormell, C. P. (1971). 'Mathematics through the imagination', *Dialogue*, **9**, 10–11.

Ormell, C. P. (1973). 'The problem of curriculum sequence in mathematics', in O'Connor, D. V. and Langford, G. (Eds.) *New Essays in the Philosophy of Education*. London: Routledge and Kegan Paul.

Ormell, C. P. (Ed.) (1974). *Polymetrics*, 7, Schools Council, Sixth Form Mathematics Project.

Peters, R. S. (1973a). 'Aims of education—a conceptual inquiry', in Peters, R. S. (Ed.) *The Philosophy of Education*. London: Oxford University Press, 11–29.

Peters, R. S. (1973b). 'The justification of education', in Peters, R. S. (Ed.) *The Philosophy of Education*. London: Oxford University Press, 239–67.

Plumb, J. H. (Ed.) (1964). *Crisis in the Humanities*. Harmondsworth: Penguin.

Pring. R. (1971). 'Bloom's taxonomy: a philosophical critique 2', *Cambridge Journal of Education*, **2**, 89.

Ryan, A. P. (1974). 'An essentially contested concept', *Times Higher Educ. Suppl.*, 120, Feb. 1st, p. 13.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Schools Council (1972). Working Paper 45, *Growth and Response*. London: Evans/Methuen.

Skemp, R. R. (1971). *The Psychology of Learning Mathematics*. Harmondsworth: Penguin.

Sockett, H. (1971). 'Bloom's taxonomy: a philosophical critique 1', *Cambridge Journal of Education*, **1**, 23.

Warnock, M. (1973). 'Towards a definition of quality in education', in Peters, R. S. (Ed.), *The Philosophy of Education*. London: Oxford University Press, 112–22.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

RESEARCH IN EDUCATION

November 1974 *Number 12*

The relation between birth order and effort or conscientiousness among primary school children: ANN START and K. B. START

Polarised cultures in comprehensive schools: W. G. QUINE

A comparison of three different measures for predicting achievement in chemistry in the age-range 15 to 18-plus: G. M. SEDDON

The preception of organisational purposes: a multivariate study of fourth-year junior school children: C. MURRAY

Comprehensive school teachers' perceptions of role-change: C. DODD

Pupils' expectations for their teachers: ROY NASH

Social class and linguistic inference: A. D. EDWARDS

Published twice-yearly, May and November
£4.50 per annum, single copies £2.75

MANCHESTER UNIVERSITY PRESS

Team Teaching in Further Education

Jack Mansell, Paddington Technical College

Summary

Over a period of five years the examination results and attitudes of part-time day technician students being taught by a team teaching organization, are analysed and compared with similar students subjected to more traditional teaching modes. The overall examination results show the team teaching results to be significantly superior. But the fluctuations over the years show the need for a more detailed study before the successful components of team teaching in further education can be identified.

Background

Much of the learning that goes on in technical colleges is on a part-time day-release basis, where young men and women students are released from their work for one day a week to pursue a course of study relevant to their job-training. Some additional evening study is usually associated with this day-release study. The usual pace of this technical college teaching is highly concentrated and exists against a background of internal or external examinations at the end of each academic session. The mode of teaching that has emerged out of this environment is essentially formal: rigid timetabling—often of over long periods, teacher-centred methods emphasising the acquisition of knowledge, repetitive drill in the solution of standard problems and always a considerable time-lag in terms of teacher-student intercourse with respect to homework, laboratory-reports and other out-of-class work. Further, as less than 40 per cent of the teachers in further education have had teacher training, the model of teaching usually adopted is that by which the teachers themselves were taught; often in the technical colleges. Thus the tendency to perpetuate the formal mode described above is strong, and is altered only gradually by the relatively few trained teachers emerging from the colleges of education (technical) and by staff engaged in in-service teacher training. Small classes of between 11 and 15 students are usual. Of recent years however there has been a greater emphasis on cost-effectiveness in this educational sector and in particular the National Advisory Council on Education for Industry and Commerce 'Pilkington' reports (1966, 1968) have resulted in larger classes in some courses.

Team teaching and the role of the teacher

The basic principles of team teaching have been explained elsewhere (Trump and Baynham, 1961; Shaplin and Olds, 1964; Beggs, 1966 and Freeman, 1969) initially in terms of American experience and more latterly in terms of British practice. It is only necessary here to remind the reader that team teaching is essentially an *organizational* concept rather than a unique combination of teaching techniques.

A progressive team will, however, watch, and where appropriate, help each other teach. They will attempt a mix of formal and informal techniques and this calls for a wider range of teacher-skills; a need identified by Lovell (1967). The role of the team-leader is as complex here as in any other group situation; it would appear desirable to have a formal leader in terms of seniority and whilst it may not be necessary or

---

Acknowledgements

This is a summary of work carried out at Reading College of Technology. I would like to acknowledge with thanks, the willingness of the Reading Education Committee to allow the project to run; to the College Principal, J. Dimmick and the Heads of Departments of Engineering, P. E. Hawes and the late C. C. Richardson, for their continued support and to those various teachers in the team who were willing to expose their teaching skills to others.

possible to have the educational expert that was envisaged in some of the earlier American team teaching literature, it is desirable to have as a team-leader someone with an appreciation of the full concept of educational technology. This job-specification requires a staffing policy less closely tied to *level* of work than now applies in most technical colleges. In team teaching the team objectives, once agreed, must be adhered to. This means that any policies with respect to handouts, marking, the issue of worked solutions, etc., must be honoured. Many teachers would identify this as an abrogation of their academic freedom. A senior teacher, marking work whilst his junior colleague lectured, might consider this to be below his dignity; older teachers may find it difficult to work at the pace of younger teachers. Team teaching is thus not suitable for all teachers. Many of the findings by Moffat (1969) and Carroll (1972) relating to the further education teacher's role in team-teaching were confirmed and it perhaps follows that any team should only be manned by volunteers.

The Reading project

It was against this background that the Reading College of Technology team teaching project was commenced in 1966. It was decided to form a multi-discipline team covering initially three subjects: mathematics, electrical science and mechanical science of the first year of the ONC in Engineering. This is a technician course and normally comprises a student population above about the 60th percentile of ability. An additional project was commenced in 1968 with the first year of an HNC course: a technician-engineer course comprising a student population normally above about the 70th percentile of ability. The comments in this report apply only to the ONC project.

The Reading team comprised five teachers with the author as team-leader, and from the preliminary discussions it was decided to apply to a day-release class the range of large-group lectures, small-group activities, and individual study periods. Certain parameters had to be adhered to: a minimum time in the science laboratories and the teacher contact-hours timetabled were not to exceed those that would have occurred with the traditional class set-up. The initial *basic* programme that emerged compared to the traditional programme is shown in Figure 1.

Figure 1: Comparison of traditional and team teaching day programmes

TRADITIONAL	TEAM TEACHING
Subject A: 2 h (electrical science)	Large group lecture Subject A: 1 h
	Large group lecture Subject B: 1 h
Subject B: 2 h (mechanical science)	Large group lecture Subject C: 1 h
	Small group Tutorial A: ¾ h
Subject C: 2 h (mathematics)	Small group Tutorial B: ¾ h
	Small group Tutorial C: ¾ h
	Individual study: ¾ h
Laboratory (A or B) 1½ h	Laboratory (A or B): 1½ h

A control group existed for each year of the project. The numbers varied from year to year according to the enrolment pattern. At no time was the class size within a control group larger than 16, which meant that two or three separate classes formed the control group over the years. The pattern of teaching for the control group was the traditional pattern described in Figure 1. The organization and timetabling for the control group was outside the control of the team teaching group. Some of the team teachers were also involved over the five-year period in teaching some of the subjects in the control groups. Where appropriate they used the same lecture materials, duplicated notes and problem solving material but always within a traditional teaching timetable and classroom allocation. The team teaching material was freely available to all teachers in the control group and it was often utilized.

The selection of the students between the team teaching and control groups was not entirely within the control of the college. Generally, students at enrolment are given a choice of a day in the week and their decisions are often a reflection of the wishes of their employers who are releasing them for study. Entrance to the course requires four appropriate 'O' levels or their equivalent or a credit performance in the General Engineering (G) Course. Attempts were made to sample the ability range of each group by the AH4 test and neither these results nor an analysis of the actual achievement levels at entry showed any marked difference between the groups. Such sampling was however far from complete.

Basic mode of operation

It must be stressed that the *order* of the lectures and tutorials for the team teaching project was not always as shown and when the educational requirements demanded it the *pattern* shown was modified: e.g. some topics such as the introduction to differential calculus, electromagnetism, gas laws, etc., were in the team's view so lengthy in terms of conceptual understanding and basic skill of application, that all the lectures and tutorials during the day were scheduled to deal with such a topic. It was reasoned by the team that this approach would produce measurable progress during the day and the students would have a better view of the goal and relative priorities. Certainly this approach tended to receive favourable comments from the students.

A study of the basic programme shown in Figure 2 will reveal the following features:

(i) The entire large group attended the three lectures. There was normally a sharing of the subject lectures and tutorial preparation between two lecturers. The limitations of the lecture method (such as identified by McLeish, 1968) were already recognized by the team and a basic rule was adopted by which all lectures were backed up with duplicated material for the students.

(ii) The number in the large group obviously depended upon the annual enrolment pattern and the physical resources available. A combination of three traditional classes each of about sixteen, giving a 'large group' of about fifty tended to emerge at Reading. In further education, a large group comprising less than three traditional classes would be of dubious value for a multi-discipline team.

(iii) The number and size of the 'small groups' will depend upon the large-group population, the size of the team and the method of timetabling the contact-hours of the teachers. In further education, the number of *class contact*

Figure 2: A typical staff and student timetable for a staff team of 5, a student population of 48, comprising 4 small groups of 72

STUDENTS				STAFF				
Group A	*Group B*	*Group C*	*Group D*	*1A*	*2A*	*3B*	*4B*	*5C*
LECTURE A				X				
LECTURE B						X		
Tutorial A	Tutorial A	Tutorial B	Tutorial B	X	X	X	X	
LECTURE C								X
Tutorial C	Lab B	Lab B	Tutorial A	X		X	X	X
Lab A	Lab	Lab	Ind. Study		X	X	X	
Lab	Tutorial B	Ind. Study	Tutorial C		X		X	X
Ind. Study	Tutorial C	Tutorial A	Lab A	X	X			X
Tutorial B	Ind. Study	Tutorial C	Lab	X		X		X

hours a teacher teaches per week is related to his grade of post and a certain number of traditional classes will consume a corresponding number of teacher-hours. Provided the number of students in a team teaching project can be equated to a suitable number of traditional classes, then a formula can be found enabling the team to be allocated a specific number of class contact hours. It is then for the team to decide how they divide those contact hours between them. For example, three traditional day-release classes such as shown in Figure 1 would require 22½ teacher-hours and if the team took on this commitment it would be a major factor in determining how frequently the students could meet in small groups supervised by a teacher. The author has shown elsewhere (1967a) that within the traditional class contact hours of technical college teachers there is opportunity for a variety of team teaching schedules. It was thus necessary for the team to enter into a 'contract' with the head of department, so that for an agreed number of contact hours the team were responsible for the education of the group within any external requirements laid down. From then on the team was responsible for the educational organization, timetabling and decision making for the students allocated to the project. One of the team's responsibilities was the structure of the small groups. As this was a developmental project, small groups ranging from seven to sixteen were tried over the five-year period and whilst it is difficult to measure the effect of this variable the team and students tended to prefer groups of twelve or less. Figure 2 shows a typical staff-student timetable for four groups of twelve. Again for any particular day, according to the decisions of the team, lectures and tutorials were interchanged and substituted for each other.

(iv) The methods used in the small-group work were decided by team discussions. In technical education much of this is guided problem-solving, but depending on the topic and associated objectives, other small-group techniques such as the discussion, seminar or structured lesson were used.

(v) The single individual-study period shown in Figure 2 is relatively modest compared to US practice and to that possible on block-release and full-time further education courses. The team was keen to include some individual study in this part-time day-release course and it was occasionally extended by shortening the large-group lectures and/or by replacing some of the laboratory periods. Various modes were tried, ranging from programmed learning, monitored worksheets, library projects—to simply offering a choice of rooms with tutors available to give guidance on any topic. The team felt that there was considerable scope for the development and evaluation of this mode of learning at the craft and technician levels in further education, and much remains to be done.

(vi) The simple model shown in Figure 2 illustrates a transient disadvantage of team teaching: at one period of the day four tutorial rooms are required for only three traditional classes. Where a larger number of small groups had been used the demand was greater. At other times of the day less than the traditional accommodation was required. Any team teaching project carried out in traditional static teaching areas comprising mainly classrooms suitable for 20-24 students, will suffer from this disadvantage and the need for more flexible teaching areas is beginning to be realized by some authorities as described by Hanson (1967) although not much as yet has been done in further education. What is required in further education at least, are teaching-learning areas that can be used for large-group teaching but be capable of being quickly partitioned to give small-group areas.

Evaluation

It was felt that if team teaching in technical colleges was to survive as an educational concept it must satisfy certain criteria: e.g.

(i) it must be acceptable to the authorities,

(ii) it must be acceptable to the students, and

(iii) it must be acceptable to the teachers.

The evaluation of the project was thus based on these criteria.

(i) An education authority will normally accept team teaching if it does not cost significantly more than methods previously used. The Reading project demonstrated that a viable team project can be mounted with no increase in teaching staff contact-time and without an *overall* increase in accommodation. The team however approached team teaching with the aim of producing a better educational effect with the same resources, rather than to produce the same effect with reduced resources. Team teaching nevertheless, may be a method of reconciling 'Pilkington' type

Table 1: Comparison of examination scores between team teaching (TT) and traditional (TRAD) control group in mathematics 01 ONC (Eng)

	Year	Mode	Number Sitting Examination N	Fail (<40%)	Pass (40–49%)	Credit (50–84%)	Distinction (85–100%)	Mean $\bar{x}$	Std. Dev. σ
1	1966/7	TRAD	41	3	4	25	9	68	18·8
		TT	38	2	7	25	4	64	16·3
2	1967/8	TRAD	8†	6	1	1	0	37	10·1
		TT	38	2	12	22	21	58	21·1
3	1968/9	TRAD	14†	2	3	7	2	60	18·9
		TT	40	3	3	31	3	62·4	15·1
4	1969/70	TRAD	14†	8	3	3	0	32·7	18·1
		TT	46	11	8	23	4	53·2	19·0
5	1970/71	TRAD	23	2	5	15	1	57·7	18·5
		TT	37	8	6	19	4	55·5	20·9
	Totals	TRAD	100	21	16	51	12	57·1	22·2
		TT	199	26	36	120	17	58·5	19·1

†These small numbers are due to some 01 mathematics students taking a different (modern maths) syllabus and thus a different examination, not included here.

criteria for course approval with small-group teaching and learning methods. Generally however, others such as Lovell (1967), Warwick (1973) maintain that the early hopes of economy from team teaching have not been fulfilled.

(ii) The students will accept team teaching if they *feel* that the method is better whilst they are involved in it, and if it gets them through the examinations no less successfully than other methods. Questionnaires were given to all those students passing through the Reading programme before they sat their examinations and the answers showed that it was generally acceptable to nearly all of them in many ways, and by a large majority each year they stated a preference to be taught again by this method. A more complete analysis of these responses has been described by the author elsewhere (1970) and similar responses have been described by Moffat (1969). The examination results over five years are shown in Tables 1-4. The overall results show that the team teaching examination results were significantly better than those of the control group.

Mathematics proved to be overall the least satisfactory subject in terms of improvement,

Table 2: Comparison of examination scores between team teaching (TT) and traditional (TRAD) control group in electrical science 01 ONC (Eng)

	Year	Mode	Number Sitting Examination N	Fail (<40%)	Pass (40–49%)	Credit (50–84%)	Distinction (85–100%)	Mean $\bar{x}$	Std. Dev. σ
1	1966/7	TRAD	44	4	7	28	5	64	17·6
		TT	38	1	3	31	3	67	13·9
2	1967/8	TRAD	29	7	7	14	1	51	20·2
		TT	39	5	4	29	1	58	16
3	1968/9	TRAD	33	9	7	14	3	51·7	22
		TT	40	5	9	24	2	58	16·5
4	1968/70	TRAD	36	26	5	5	0	32·4	16·1
		TT	46	20	9	17	0	47·0	20·5
5	1970/71	TRAD	23	8	5	9	1	47·2	23·3
		TT	37	6	9	20	2	53·1	19·6
	Totals	TRAD	165	54	31	70	10	50·0	22·5
		TT	200	37	34	121	8	56·3	18·8

Table 3: Comparison of examination scores between team teaching (TT) and traditional (TRAD) control group in mechanical science 01 ONC (Eng)

	Year	Mode	Number Sitting Examination N	Fail (<40%)	Pass (40–49%)	Credit (50–84%)	Distinction (85–100%)	Mean $\bar{x}$	Std. Dev. σ
1	1966/7	TRAD	41	5	11	24	1	56	15·8
		TT	38	4	5	23	6	53	19·7
2	1967/8	TRAD	26	7	6	12	1	50	17·6
		TT	39	7	6	25	1	55	19·3
3	1968/9	TRAD	31	9	7	12	3	51·7	22·5
		TT	40	6	4	29	1	56·5	20·9
4	1969/70	TRAD	33	12	6	13	2	45·5	21·7
		TT	46	10	9	22	5	55·9	21·5
5	1970/71	TRAD	24	5	3	13	3	56·2	24·2
		TT	37	2	9	25	1	59·1	15·9
	Totals	TRAD	155	38	33	74	10	51·9	20·7
		TT	200	29	33	124	14	55·9	19·8

and in view of the large amount of individual exercise and tuition traditionally associated with this subject this is perhaps not surprising. This was on the other hand the subject that had the most staff changes over the five year period, and for two years the team teaching results were significantly better than those for the control group.

Electrical Science was the most satisfactory in terms of examination results and although in one year only were the results significantly better, in all years they were never lower than those of the control group. The staff in this subject remained the same over the full five year period. The subject is considered by many teachers to be at least as conceptually demanding and requiring as much individual tuition as mathematics.

The mechanical science team teaching examination results were only significantly better than the control results in one year. This is a subject not unlike that of electrical science in terms of the new body of technological experience to be assimilated by the students but not quite so mathematical. Some changes in teachers did take place in this subject during the last year of the project.

The examination wastage rate in percentage terms was, for the team teaching group, considerably lower than for the control group. Other wastage occurred during the academic session by students leaving the course. Again an attempt was made to sample this by examining in detail the reasons for some of the students leaving. Generally this type of wastage fluctuated between eight per cent and 20 per cent over the years and the reasons included academic inability to cope with the demands of the course, change of job and/or address, or for no apparent reason. The pattern of wastage appeared to be about the same for both groups.

Generally the team teaching examination results compared worst at the end of the first year and went on to improve in successive years, although the improvement was by no means uniform or consistent. The pattern is similar for all three subjects and this may reflect either the general ability levels of the group each year and/or the staffing continuity.

(iii) The acceptibility of this mode of teaching to the lecturers can be indicated by the degree of willingness to remain in the team. They were all volunteers and each session each lecturer had the option of withdrawing if he so wished. Of the original five members comprising the team, four of them continued until well into the final academic session when two of them left to take up posts in other colleges. One other original member of the team had a year out of the team in the 1967/68 session due to other departmental duties and withdrew from the team after the 1969/70 session. His replacement in 1967/8 was pleased to come out of the team at the end of that session as he felt that his style and organization was not compatible with that of the team. The other replacement in 1970/71 would have been pleased to continue had the project continued as he found working in the team suited him. This

Table 4: Comparison of aggregate examination results between team teaching (TT) and traditional (TRAD) control group over five years

MODE	N	FAIL (<40%)	PASS (41–49%)	CREDIT (50–84%)	DISTINCTION (85–100%)	MEAN $\bar{x}$	STD. DEV. σ	LEVEL OF SIGNIFICANCE ρ
MATHEMATICS								
TRAD	100	21	16	51	12	57·1	22·2	
	(100%)	(21%)	(16%)	(51%)	(12%)			
TT	199	26	36	120	17	58·5	19·1	NS
	(100%)	(13·1%)	(18·1%)	(60·3%)	(8·5%)			
MECHANICAL SCIENCE								
TRAD	155	38	33	74	10	51·9	20·7	
	(100%)	(24·5%)	(21·3%)	(47·7%)	(6·5%)			
TT	200	29	33	124	14	55·9	19·8	NS
	(100%)	(14·5%)	(16·5%)	(62%)	(7%)			
ELECTRICAL SCIENCE								
TRAD	165	54	31	70	10	50·0	22·5	$p<$ 0·05
	(100%)	(32·7%)	(18·8%)	(42·5%)	(6%)			
TT	200	37	34	121	8	56·3	18·8	
	(100%)	(18·5%)	(17%)	(60·5%)	(4%)			
TOTALS								
TRAD	420	113	80	195	32	52·4	21·9	$p<$ 0·01
	(100%)	(26·9%)	(19·1%)	(46·4%)	(7·6%)			
TT	599	92	103	365	39	56·9	19·2	
	(100%)	(15·3%)	(17·2%)	(61%)	(6·5%)			

general pattern of continuity appears to reflect a degree of commitment of the staff in personal terms to this particular mode of team teaching and it can be concluded that team teaching was acceptable to the majority of teachers on this project over this period.

Discussion

As this was an operational project, not all the variables were under control. The size and nature of the control groups varied from year to year, as did the staff teaching them. The technical syllabuses were modified during the project and both the team teaching and the control groups had to cope with a common problem of reduced time allowance for these syllabuses. One consequence of this modification was the inclusion of a General Studies section into the team during the last two academic sessions. The role of affective learning in technical education has only recently been given its due weight and even now the teachers of general studies are often faced with considerable antagonism. A large-group lecture plus small-group discussion mode was set up within the timetable and the results, although difficult to evaluate were encouraging. That it is possible to bring about affective learning by large-group lectures plus small-group discussions has been shown by Hoover (1969). Team teaching offers an opportunity to discover the optimum proportions of this combination.

The examination results tend to suggest a growing expertise by the team in the handling of the educational strategies and teaching methods called for by this form of team teaching. Further work needs to be carried out before it is possible to discern whether the better overall results were due to simply this growing expertise of the team or due to some other more intrinsic property of team teaching. The fact that some of the team teachers were often engaged in teaching the same subjects to the control groups tends to support the concept of team teaching as a successful innovation. The fewer examination failures may have been due to the better monitoring of student performance by the team. The results also tend to show that it is the average student who appears to benefit the most from team teaching and this has been suggested elsewhere by Beggs (1966) and others. It was generally recognized by the team, particularly in the early years, that they were working much harder both in and out of the classroom, than they would have done on a more traditional method. It appears that the level of preparation, response of marking, keeping of records, etc. agreed by the team were deliberately held at a high level by the rules of the team initially agreed. It would follow from this that newcomers to the team would not have this degree of commitment. More work needs to be done therefore not only on whether this is a general result of team teaching but also on how long and under what conditions this extra activity is maintained. The falling results in the final year could have been a result of staff leaving but it may have been a running out of innovative steam.

All the team members took some time to adjust to the large group lecture technique and if team teaching is to expand in further education, the colleges of education (technical) may have to give more guidance in this instructional mode.

In this experiment it has been shown that team teaching can be effective for and acceptable to technician and higher-technician grade students, but there remains much work to be done in evaluating its successful components.

References

Beggs, D. W. (Ed.) (1966). *Team Teaching*. Indiana University Press.

Carroll, A. D. (1972). 'Team teaching in further education', *Education and Training*, August.

Freeman, J. (1969). *Team Teaching in Britain*. London: Ward Lock Educational.

Hanson, W. J. (1967). 'Teaching areas for team teaching', *Higher Educational Journal*, Spring.

Hoover, K. H. (1969). 'Flexible scheduling and affective learning', *Educ. Res.*, **12**, 1, 70-3.

Lovell, K. (1967). *Team Teaching*. Leeds University, Institute of Education.

Mansell, J. (1970). 'Team teaching—an analysis of students' views', *The Technical Journal*, February.

Mansell, J. (1967a). 'Team teaching in further education', *Technical Education*, June.

Mansell, J. (1967b). 'Team teaching and conditions of service', *The Technical Journal*, October.

McLeish, J. (1968). *The Lecture Method*. Cambridge Institute of Education.

Moffat, T. (1969). *Team Teaching*. Occasional Paper, Durham Technical College.

National Advisory Council on Education for Industry and Commerce (1966). Report of Committee on Technical College Resources. *The Size of Classes and Approval of Further Education Courses.*

National Advisory Council on Education for Industry and Commerce (1968). Report of Committee on Technical College Resources. *The Use of Buildings and Equipment*.

Shaplin, J. T., and Olds, J. F. (Eds.) (1964). *Team Teaching*. New York: Harper and Row.

Trump, J. L., and Baynham, D. (1961). *Focus on Change—Guide to Better Schools*. Chicago: Rand McNally.

Warwick, D. (1973). 'Where method meets theory', *Times Educ. Suppl.*, 30th March.

Interdisciplinary Education

Lowering the Barriers to Effective Learning

John P. Wilson, Boston University School of Law
Val D. MacMurray, Faculty of Medecine, University of Calgary Medical School

Summary

The structure of modern, graduate level education has impeded the development of interdisciplinary programmes which emphasize breadth rather than depth of inquiry. Yet societal problems frequently require interdisciplinary resolution and an understanding of the values and assumptions of the representatives of different disciplines. An experimental programme at Boston University, involving the exploration of existing problems in the health field by small, interdisciplinary groups of graduate and professional students, is an attempt to meet the need for interdisciplinary education. An analysis of two of these groups has demonstrated that they must resolve problems of *acceptance*, *data-flow*, *goal-formation* and *control* before they can function effectively as an innovative and useful departure in graduate level education. In the evaluation described in this article, it was found that for significant group development to occur, each dimension had to be successfully mastered.

'Before I built a wall I'd ask to know
What I was walling in or walling out'

Mending Wall
Robert Frost

Introduction

The increasing complexity of society which creates a need for new areas of specialization, has also expanded the number of problems requiring a multidisciplinary resolution and has created a new need for persons who can co-ordinate and harmonize different disciplinary approaches. Particularly in the resolution of complex 'social' problems, it is recognized that knowledge from several areas is necessary to achieve adequate, well-balanced, and long-lasting solutions (Forrester, 1971). An emphasis upon one theoretical approach or academic discipline without taking account of others is likely to produce results more deleterious than the situation which created the original problem.

Yet within our universities, the growth of knowledge, the need to produce specialists, and the requirements of increasingly narrow and exact research within each area of knowledge have led to the creation of different schools and departments within schools. As the growth of knowledge within each field has come to exceed the grasp of the most intelligent scholars, subfields have proliferated, e.g. urban sociology, Asian history, labour economics, corporate law; and mastering even one has become increasingly difficult. This development has split the community of educated men so that they comprehend neither the discourse nor the values of each other, to the detriment of all.

To remedy this situation, several people within and without educational institutions have urged change in our highly differentiated academic structure. Henry Winthrop (1959, 1965, 1966, 1967), for example, in considering interdisciplinary collaboration, stresses the 'spirit' of any interdisciplinary effort which places primary value on the process of intellectual integration. Similarly, Wise (1966), in elaborating on the values of integrative education in a disintegrated world, argues for an effective exploration of a 'post-modern' multiverse to replace the current irrelevance of compartmentalized curricula in our colleges and universities[1]. Discussing the inhibiting nature of 'categories', Wise states that, like the modern ideologies which often reflect the provincial borders of our intellectual disciplines by reducing man to *only* his economic or *only* his psychological or *only* his technological or *only* his political dimensions, these 'categories' have shielded us from the rich texture of existential reality.

---

[1] Our world might be termed *post modern* in that the disintegrating nature of the great consensus of progressive rationalism characterizes the epoch beginning in the Renaissance and extends to World War II.

Multiverse refers to a disintegrated world blown apart by expanding knowledge and war yet thrown together by technology and science; a world where intellectual unity can no longer be assumed nor can the combining of disciplines simply add up to an integrating or liberating education for either undergraduate or graduate students.

An equally critical view has been expressed by William Arrowsmith (1969), who has made the following assessment of the quality of graduate education:

> There is incoherence in graduate educational programmes, a disturbing trend toward conformity of curricula, and a severe neglect of the educational needs and aspirations of graduate students themselves. It is practically impossible to get a group of aggressive specialists to agree on any special programme.

In reference to professionalization, Arrowsmith argues that its typical dangers are disciplinary rigidity, narrow-mindedness, and resistance to change. One of the worst features of graduate education is its failure to grant students initiative and freedom, thereby effectively blocking, rather than stimulating, intellectual curiosity. Rare, indeed, says Arrowsmith, is the graduate department which encourages a student to look beyond his own field.

The development of interdisciplinary education

Increasingly, academic institutions are experimenting with the concept of multidisciplinary offerings. Now clusters of combinations of courses have begun to emerge such as systems engineering, urban studies and regional studies. All of these new approaches, in varying degree, focus less on the mastery of a particular discipline and more on the application of several disciplines to the resolution of specific problems.

Two research psychologists at the Center for Research and Development in Higher Education at the University of California, Berkeley, have reported recently that interdisciplinary study has become an increasingly important trend in American education. They indicate that a number of new courses which focus upon social problems, intellectual themes, or human experiences have been introduced at hundreds of schools (Brick and McGrath, 1969); ethnic studies programmes are increasingly common (Wise, 1966); and many cluster colleges offer alternative liberal arts curricula which are frequently interdisciplinary (Gaff, 1969). For example, in the fall of 1973, Tufts University announced an undergraduate interdisciplinary curriculum enabling students in a variety of fields to pursue an organized programme in Community Health and the Delivery of Health Services. While the programme has its focus in the colleges of Liberal Arts and Engineering, the curriculum also draws upon the resources of its professional schools, in particular the Schools of Medicine and Dental Medicine. The programme is intended to acquaint students with social, economic, political, psychological, and other aspects of the health care delivery system. Moreover, at Tufts and many other Schools, students are increasingly permitted, under 'independent study' arrangements, to construct their own courses or even entire programmes from materials and courses not offered by any single department.

There has been little documentation and systematic evaluation of interdisciplinary approaches in our educational institutions, but the extant literature indicates that the interdisciplinary approach has been employed in a variety of ways with a variety of disciplinary combinations. (See Sargent and Smith, 1949; Carroll, 1951; Anderson, 1959; and Luszki, 1958.) Although many of these efforts may have been effective, poor reporting has left us with only scanty impressions of their educational effectiveness and little information regarding the problems or processes unique to interdisciplinary learning.

This dearth of 'hard' information has fueled the charges of those educators who maintain that interdisciplinary work results largely in superficial breadth without a genuine understanding of the knowledge and approaches of each contributing field. This attitude is mentioned by Wise (1966), who states:

> The attempt to relate together entire fields does not help us integrate in a disintegrated world. For if mastering one discipline is now impossible, how much more precarious it is to pretend knowledge of two or three. Such an interdisciplinary attitude is bound to meet the charge of dilettantism, a charge which in this context seems largely justified.

It is difficult to refute this charge, and, indeed, it is likely that in many instances it cannot be refuted. But it may be answered by an assertion that some breadth, rather than depth, is valuable for its own sake in a narrowly confining curriculum. It may be further answered by an approach which does not require students 'to relate together entire fields'. Complex problems often require interdisciplinary analysis, students may learn significantly about the knowledge and methodology of each contributing discipline in regard to the specific problem, and they may absorb much of genuine value about the processes of interdisciplinary collaboration.

An evaluation report from Boston University

The Center for Law and Health Sciences at Boston University School of Law is an organization which attempts to implement a 'problem solving' programme in interdisciplinary education. It was established with a multi-disciplinary faculty and student body drawn from several educational institutions in the Boston area. The Center's programmes focus upon different ramifications of one major aspect of society—health. Its educational philosophy maintains that innovative, interdisciplinary programmmes are necessary, both to assist in the resolution of current problems and to train professionals in different fields about the knowledge and approaches of other disciplines and the techniques by which interdisciplinary analysis is performed.

This paper analyses one of the Center's educational programmes: interdisciplinary education/research by collaborating faculties and students organized into teams of approximately six members. There are several basic assumptions underlying the design of this programme:

(1) Interdisciplinary education should supplement, not replace the work of students in their separate disciplines.

(2) Its programmes should primarily subject students to a value confrontation, i.e., the need to reconcile the value judgements made by students from several disciplines.

(3) Value confrontation requires groups small enough to engage in free, frequent interchange of thoughts and opinions.

(4) Value confrontation is facilitated through joint analysis and attempted resolution of real problems of immediate concern to society. The 'reality' of the problem, therefore, requires 'reality' in the process of resolving it, forcing students to assume unfamiliar roles which many may occupy in later life.

(5) The problems selected for analysis should be carefully circumscribed.

(6) By requiring a written, carefully designed research product, an essentially unstructured, superficial experience without direction or goal is avoided. The requirement of a product enhances the quality and depth of the process—that is, the student synthesis of values and learning and subsequent broadened perspective that is the goal of interdisciplinary education. It also provides a means to evaluate the educational process.

Preliminary observations of learning effectiveness

The team research programme of the Center has generated substantial enthusiasm and interest from student participants. Not all groups, however, have enjoyed effective learning experiences, and subsequent evaluation indicates that the breakdown of interdisciplinary collaboration was related to group development and maturity. Our experiences indicate that interdisciplinary education most frequently fails because its methodology and design repeatedly ignore problems and processes of human social action, interaction and group development.

Luszki (1958) explicated these problems of close collaboration in group functioning, and these considerations remain the paramount focus of any concern in any interdisciplinary group effort. The major failing of interdisciplinary research or education is not the inability to blend or mix divergent methodologies and bodies of knowledge, but rather the inability to mix human beings in group situations which require, even demand, a human synthesis to facilitate group development. Only to the extent that a group is able to resolve particular intra-group problems or concerns, ubiquitous to groups generally, will it be able to develop intellectual synthesis.

Gaff and Wilson (1971), discussing faculty interaction in an interdisciplinary setting, confirm this finding. They argue that:

> To the extent that innovators find ways to effect a kind of cultural pluralism—that is, are able to create a climate of respect for alternative habits of mind and style, in which differences are considered complementary rather than in conflict—they can be expected to create interdisciplinary programmes that are not only successful but immeasurably enriched.

The crucial question for the innovator, for the organizer of an interdisciplinary experience, is *how* 'to create a climate of respect for alternative habits of *mind* and *style*'. In an attempt to answer this question, observations of the experiences of two groups will be discussed. Both groups were composed of graduate and professional school students in their early to mid-twenties. The membership of each group, and the problem each confronted, were as follows:

Group A. This group considered features of present and alternative health care delivery systems and the impact upon them of third party payment mechanisms. It was composed of six members, five of whom were active (one each in law, social planning, political science, economics, and urban studies) and two faculty advisers (one in medicine and one in economics).

Group B. This group explored issues related to the ethical and legal implications of recent genetic research. It comprised five team members (one from a divinity school, one from a law school, two social planners, one bio-engineer) and two faculty advisers (one in social planning and one in medicine).

In analysing the development (maturity) of each group and its capacity to facilitate personal learning, a conceptual model was employed which is derived from research on a wide variety of groups in various life settings. This model indicates the apparent validity of the assumption that there are four basic dimensions of personal and group growth that arise inevitably from all social interaction (Bradford *et al.*, 1967). These dimensions are: Acceptance, Data-Flow, Goal-Formation, and Control. They are in no way discrete but in fact are highly interdependent.

Acceptance

Acceptance is related to the formation of trust, the amount of acceptance the members have for themselves and for others, the reduction of fear of self and of others, and consequent feelings of confidence. Acceptance, in short, is related to a mutual 'supportive climate' within a group, and acceptance seemed to prevail particularly in Group A. The members in this group were very supportive of one another during those occasions when feelings surfaced and intellectual arguments took the form of personal attacks. Usually, at these times, several members intervened to assist or to initiate processing some of the personal issues that emerged. On the other hand, Group B was characterized by a more cool and disengaged atmosphere. There was less sharing of personal feelings and experiences, suggesting that the group had not established the clear and cohesive boundaries of Group A.

Ultimately, questions of acceptance and the formation of trust become questions of membership and inclusion. In any group, certain individuals, for either personal or intellectual reasons, or both combined, achieve dominance. Their assumption of this role, however, in the absence of a clear designation of authority, must be based upon the tacit approval and consent of the other members. If this is not achieved, or is achieved only partially, then internal dissidence, rather than harmony, is likely to characterize the group's working relationships. Moreover, if one or more group members are unwilling to modify or abandon their own ideas, or are granted overt preference by academic teaching staff who are sympathetic to their ideas and intellectual styles, yet the majority have competing goals, objectives and values, then the former become 'outsiders', performing their work apart from, and at variance with, the majority of the group. Only firm external controls, or the presence of a designated team leader vested with authority to resolve disputes, can avoid this result.

In the groups under discussion, although one developed more acceptance and trust than the other, both acquired forces or processes which substantially reduced concern about acceptance and thereby permitted concern about problems of communication. This is the second major dimension of group development.

Data-flow

Data-flow relates to the flow of attitudes, feelings, and perceptions between the members of a group. Ultimately it is expressed by decision-making and the selection of alternative courses of action.

Because of the general climate of trust (although in different degrees) in both groups, there was a fairly free flow of communication about the way the group members felt, their perceptions, and their attitudes in relation to relevant concerns of the group. Often in groups so much emphasis is placed upon intellectual problem-solving that a tendency arises to refrain from dealing with feelings and perceptions. Rather than express feelings, members may deny their existence and develop skill in disguising them. Efficiency is reduced, however, when defensive rather than supportive climates exist.

To a considerable extent, both Group A and Group B developed procedures which helped them to listen, to express feelings and perceptions with candor and honesty, to be spontaneous, and to integrate (with the substantive material) emotionality into their work. As expression such as, 'I agree with your point, but that makes me mad' was not uncommon. Group members exchanged feelings nearly as freely as they exchanged ideas related to their substantive problems, but the partial blocking of inter-personal exchange in Group B also partially blocked intellectual exchange and its problem-solving, decision-making abilities.

Data-flow was further increased in Group A by the absence of a language barrier—or, expressed inversely, by the presence of a 'language community'. Even though there were students representing economics, law and urban sociology in the group, there was little reliance upon the jargon or specialized vocabulary of each discipline. In their speech, the students could not be readily differentiated by their disciplinary backgrounds. As might be expected, symbol sharing and the bridging of the different language communities was a greater problem in Group B. There members relied more on specialized vocabulary, at times preventing clear articulation of problems and cross-fertilization of ideas.

Goal-formation

Goal formation relates to the motivation of individual team members and the concerted implementation of these motivations by the total group. It is the joint selection and accomplishment of a designated task.

Although both groups were generally successful in developing acceptance and data flow, questions concerning group productivity and the formation of joint goals were, to a large extent, unresolved. The members continually queried the purposes and value of producing a joint report. Their efforts in this direction were not well co-ordinated and arose more from a sense of duty than from a sense of personal commitment.

Possibly the failure to master this aspect of group development stemmed from the short duration of each group's effective working life (one semester). There was insufficient time to master the first two dimensions of group development, i.e., to create an interdisciplinary interchange *and then* to decide upon the goal of an interdisciplinary product. Certainly the restraint of time and the novelty and difficulty of producing a joint report synthesizing the group's work were factors which frustrated goal formation. But there was more. Both groups had difficulty even accepting the objective of a joint report, possibly because this goal did not evolve out of the group's interaction but out of a programme structure which appeared to be external to and separate from them. Another, more practical reason may have also influenced each group's attitude: Because many students were required to write individual papers to satisfy credit requirements at their schools of registration, the joint report was all too easily viewed as an irksome and superfluous burden.

For whatever reason, the concept of 'product' enhancing 'process' did not materialize. Product and process were viewed as dichotomies; valuable time was sacrificed to discussion concerning the best way to accomplish the predetermined goal, with a resulting diminution of discussion and argument concerning substantive issues.

Control

This dimension concerns the mechanisms, e.g., the democratic or authoritarian process of appointing a leader, by which tasks are assigned in the accomplishment of a common goal. It can be described as those regulatory devices that lead to co-ordinated sequences of behaviour in a group and the formation of roles and expectations. Eventually this process concerns organization—the degrees of formality, stability, and complexity in all varieties of social relationships.

This concern, like the last, was only partially resolved by the two groups—often for the same reasons. Most strategies for leadership and control developed around the task of producing a joint report, and as neither group had a clearly recognized leader, the variety of persuasive methods members used for controlling others, e.g., 'advice giving', 'debate and argument over procedure' (not content), did not resolve the problem. Because confusion about 'who's in charge' and 'do we really want to write a report' appeared to threaten trust and openness of communication, power struggles and conflicts were not openly worked through. Thus, allocation of work occurred more through bartering ('I'll do this part if you'll do that') than through interdependent formation of roles or allocation of roles by a person with designated authority.

Although the issue of control raised a number of problems, it posed an opportunity for potential learning. Such learning did not take place because of the hesitance of both groups to deal adequately with their own internal processes at those moments when discussion centred on producing a report. At those times the groups tended to regress in terms of their capacity and ability to process their experience and focused instead on group maintenance.

On the other hand, when people were doing what they wanted to do (discussion), knew that they were doing what they wanted to do, and had trust and acceptance of each other in the situation, the control and organization problem became relatively simple or disappeared. Problem-solving roles were easily distributed and were even allocated spontaneously in response to immediate problem-solving needs. At these times conflict seemed to disappear, and the groups were productive and directed toward problem solution.

Conclusions

Four dimensions of group interaction were identified in observation of the two student groups: acceptance, data-flow, goal-formation, and social control. For growth to occur, each dimension had to be successfully mastered. Significant reduction of the acceptance and data-flow problems occurred, but the external demand of 'producing a product' raised a number of issues for both groups which precluded attainment of goal-formation and adequate control. This is not to say that the groups were unproductive—they achieved most of the objectives set for them and developed sufficient cohesion and trust to facilitate an interpersonal, intellectual exchange; all members of both teams submitted final reports, but they were individual reports rather than a synthesized 'product' of each team's work.

In order to deal with the problems which were identified, the following changes were mandates:

(1) *Orientation.* Establishment of a brief orientation programme, prior to inception of the actual group experience, to acquaint students with the nature of group process and the skills and methods required to deal with the predictable developmental problems of groups.

(2) *Structure.* Appointment of a student leader with clearly identified status to provide leadership and control within the group yet maintain a peer relationship with other members to facilitate personal and intellectual interaction; faculty should continue to act as informal advisers.

(3) *Product.* Either attainable goals should be formulated for groups in advance, or, preferably, groups should determine their own common goals which might include joint reports or significant alternatives, e.g., draft legislation or regulations, symposium, design of experimental project.

Interdisciplinary student teams, structured according to the model described above, represent an innovative and significant departure in graduate level education. The design adopted at Boston University is susceptible of easy replication elsewhere. These teams are a broadening experience in an educational system which has narrowed the range of subjects available to an advanced student.

Participating faculties also benefit from the experience. Most are accustomed—although perhaps not without substantial reservation—to the guild structure of American higher education, to the graduate schools and departments which have artificially divided the recorded sum of human experience. Their participation is a recognition of the need to integrate subjects once made vigorously distinct. And through their participation, and the development of strong interdisciplinary programmes, the structure of education may be modified to meet the demands and needs of modern society.

References

ANDERSON, H. H. (1959). *Creativity and its Cultivation.* Addresses presented at the Interdisciplinary Symposia on Creativity 1957-8, Michigan State University. New York: Harper and Row.

ARROWSMITH, W. (1969). 'Innovation and reform in graduate education', *Boston University Journal,* **XVII,** 2 and 3.

BRADFORD, L. P., BIGG, J. R. and BENNE, K. D. (1967). *T-Group Theory and Laboratory Method.* New York: J. Wiley and Sons.

BRICK, M., and MCGRATH, E. J. (1969). *Innovation in Liberal Arts Colleges.* New York, Teachers College, Columbia University.

CARROLL, J. B. (1951). Interdisciplinary Summer Seminar in Psychology and Linguistics, Cornell University.

FORRESTER, J. W. (1971). 'Counterintuitive behaviour of social systems', *Zero Population Growth National Reporter*, June.

GAFF, J. G. (1969). 'Innovation and education: a case study', *Education Record*, **L**, Summer, 290-9.

GAFF, J. G., and WILSON, R. C. (1971). 'Faculty cultures and interdisciplinary studies', *Journal of Higher Education*, **XII**, March, 186-201.

LUSZKI, M. E. B. (1958). *Interdisciplinary Team Research: Methods and Problems*. Published for National Training Laboratories by New York University Press.

SARGENT, S. S., and SMITH, M. (Eds.) (1949). *Culture and Personality*. New York: Wenner-Gren Foundation for Anthropological Research.

WINTHROP, H. (1959). 'Some values of intellectual synthesis', *Social Studies*, **L**, December, 243-50.

WINTHROP, H. (1965). 'Contemporary intellectual ferment and the curriculum of the future', *Social Studies*, **LVI**, March, 83-9.

WINTHROP, H. (1966). 'An institute of intellectual synthesis', *Education Forum*, **XXX**, January, 221-9.

WINTHROP, H. (1967). 'Specialization and intellectual integration in liberal education', *Educational Theory*, **XVII**, January, 25-31.

WISE, G. (1966). 'Integrative education for a disintegrated world', *Teachers College Record*, **LXVII**, March, 391-401.

British Journal of Educational Psychology

Prof. N. J. ENTWISTLE, Department of Educational Research, Cartmel College, Lancaster

Vol. 44 Part 3 November 1974

J. D. NISBET: Editing the Journal.
M. H. NEVILLE and A. K. PUGH: Context in Reading and Listening: a comparison of Children's Errors in Cloze Tests.
G. W. HERBERT: Teachers Ratings of Classroom Behaviour: Factorial Structure.
A. G. SMITHERS and S. DANN: Success and Failure among Engineers, Physical Scientists and Linguists at a Technological University.
H. PHILP and M. KELLY: Product and Process in Cognitive Development: Some Comparative Data on the Performance of School Age Children in Different Cultures.
S. C. SOMERVILLE: The Pendulum Problem: Patterns of Performance Defining Developmental Stages.
R. D. DUTCH and J. MCCALL: Effects of Introducing a Transition Department in a Secondary School.
H. FRANCIS: Social Background, Speech and Learning to Read.

RESEARCH NOTES:
R. J. RIDING and J. M. SHORE: A Comparison of Two Methods of Improving Prose Comprehension in Educationally Sub-Normal Children.
A. G. SMITHERS and D. CHILD: Convergers and Divergers: Different Forms of Neuroticism?
P. M. MILLER and R. R. DALE: A Comparison of the Degree Results of University Students from Co-educational and Single-sex Schools.
C. EISER: Recognition and Inference in the Co-ordination of Perspectives.
D. BOYDELL: Teacher-Pupil Contact in Junior Classrooms.

BOOK REVIEWS.

PUBLICATIONS RECEIVED.

INDEX TO VOL. 44.

NOTES TO CONTRIBUTORS.

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Limitations of Personality Testing

as Applied to College of Education Students

F. C. Lamb, late of St Osyth's College, Essex

Summary

A whole year intake of a college of education was tested near the beginning and end of their course using two tests, the Study of Values and the Minnesota Teacher Attitude Inventory. Results were analysed in groups by testing changes in mean scores for significance. Few significant changes were found over a period of two-and-a-half years. An examination of the tests and their bases leads to some doubts as to the validity of the tests in their present form.

Introduction

Evans (1967) attempted to evaluate some of the changes in the values and attitudes of a group of post-graduate students during their time in a university department of education. She suggested that if some values and attitudes are desirable in teachers and these do not change, selection of students must be made on the basis of these values and attitudes rather than on the basis of those which do change. The group consisted of 46 women and 32 men who had graduated in June, 1966, their ages at graduation being not over 23. The students were tested in October, 1965 and again in May, 1966. The tests used were Study of Values, British Edition (Richardson, 1965), and the Minnesota Teacher Attitude Inventory (Cook *et al.*, 1951).

With Study of Values in October, 1965, significant differences between the scores of men and women were found in all values except the religious value. In May, 1966 the differences between the religious and the economic values were not significant. With these exceptions all other differences were significant at the five per cent or one per cent level. So it was argued that men and women must be treated as two different populations. On the whole there were no marked differences in scores over the period for either men or women. Evans suggested that for a longer course and with younger students larger differences might show. In the case of the Minnesota Inventory, women scored significantly higher than men on both occasions, and scores of both men and women rose significantly on the second testing.

The tests

The Study of Values test is based on Spranger's work 'Lebensformen', quoted by Allport (1963), in which Spranger made a case for judging the personalities of men by their valuative attitudes. He distinguished six clusters, all of an idealistic type, and allowed that most men are of mixed types. The chief attributes of the six types are as follows:

1. *Theoretical.* The main goal is the search for truth. By thought, experiment and criticism the theoretical man strives to order and systematize knowledge.
2. *Economic.* Usefulness is the highest value, and business the best activity. Wealth and the goods of this life are eagerly sought after.
3. *Aesthetic.* Form and harmony are the finest ends. Beauty is considered superior to truth.
4. *Social.* Love of people is the highest value. At best the social type is selfless and approaches the religious type.
5. *Political.* The chief interest is power, not merely political power but also control of other people.
6. *Religious.* The unity of the whole universe is what the religious man works for, and this goal is often sought for by self-denial and asceticism.

The Study of Values test was originally devised by G. W. Allport but the version used was the British Edition of 1965 revised by Sylvia Richardson. The test consists of a booklet in two parts. In Part 1 the testee is given 30 alternative choices which can be resolved by allocating scores of 0, 1, 2, or 3 to one alternative and 3, 2, 1, or 0 respectively to the other. In part 2 there are fifteen items and in each item four implied values are ranked in order of preference using scores of 3, 2, 1, and 0. Each value occurs the same number of times as all the others in both parts and the total score for each part is fixed so giving a quick rough check on the scoring.

The alternatives offered in Part 1 are similar to the following: a preference for writing as against engineering would indicate an aesthetic value rather than an economic one; joining a local political party group rather than attending church would indicate a political value as against a religious one; rating Faraday as more important than Elizabeth Fry indicates a theoretical value while the converse shows a social value.

The sort of choices offered in Part 2 are exemplified by the following: participating in a social club, in an electoral campaign, in a choir and in a religious revival have respectively social, political, aesthetic and religious values; assessing the value of science as a system of knowledge, as an extension of man's control over nature, as a widening of the understanding of man and as a deepening of man's wonder at the marvels of the universe would be associated with the values of a theoretical, economic, social and religious type respectively.

Scores for each value are totalled and the result plotted on a graph on the back of the booklet. For comparison the norm is printed on the graph so that the individual can compare his profile with that of the norm.

The structure of the test forces the testee to compare two values at a time in Part 1 and four at a time in Part 2. So it is the relative strengths of the values that are measured, not their actual strengths. Thus an energetic, dynamic person could produce the same profile as an apathetic, lethargic individual, while at the same time the weakest value of the former could well be, in absolute terms, much stronger than the strongest value of the latter.

The Minnesota Teacher Attitude Inventory is clearly based on the composers' concept of the ideal teacher. As the manual states,

> It is assumed that a teacher ranking at the high end of the scale should be able to maintain a state of harmonious relations with his pupils characterized by mutual affection and sympathetic understanding. The pupils should like the teacher and enjoy school work. The teacher should like the children and enjoy teaching. Situations requiring disciplinary action should rarely occur. The teacher and pupils should work together in a social atmosphere of co-operative endeavour, of intense interest in the work of the day, and with a feeling of security growing from the permissive atmosphere of freedom to think, act and speak one's mind with mutual respect for the feelings, rights and abilities of others. Inadequacies and shortcomings in both teacher and pupils should be admitted frankly as something to be overcome, not ridiculed. Abilities and strengths should be recognized and used to the utmost for the benefit of the group. A sense of proportion involving humour, justice and honesty is essential. Group solidarity resulting from common goals, common understanding, common effort, common difficulties, and common achievements should characterize the class.

The test is composed of 150 short statements about children, adults and teachers and classroom situations. Some of these statements recur in a more or less disguised form and some testees realized that they were being asked the same question more than once. The testee is asked to indicate on a separate sheet whether he strongly agrees, agrees, is uncertain, disagrees or strongly disagrees, giving a five-point scale for each item.

From the point of view of the test constructors some responses are 'right', i.e. in accordance with their views, and some are 'wrong'. 'Right' answers are positive and 'wrong' answers negative, the score being the algebraic sum of the two part scores, which are easily obtained by placing first the 'rights' key on the answer sheet and counting the answers that show through the holes in the key, and then repeating with the 'wrongs' key.

The investigation

It was decided to act on Evans's suggestion so the 1967 intake at the college in which the experiment was conducted was given Study of Values and Minnesota Teacher Attitude Inventory as soon as possible. The intake consisted of men and women and they were divided into the groups usually then found in colleges of education, viz. infant, infant-junior, junior, and junior-secondary. While there were women in all these groups, there were no men in the infant groups. The numbers of men in the groups increased on ascending the age range so that in the junior-secondary groups the men outnumbered the women. In addition a small group of women were taking a specialist needlecraft course and a larger group taking a home economics course. These last two groups were not included in the first testing and so were excluded from the main analysis which dealt with those students who completed both tests on both occasions. There was a considerable wastage for a variety of reasons so that the final population was 112 women and 32 men. Results from the 43 specialist students were used in an extended analysis.

In May, 1968, when the tests were first administered, the students had completed one teaching practice, the first of three. The great majority were 18-plus years of age on entering college in September, 1967. A few were older, but none were over 25 years. The second test was adminstered after the final teaching practice, and before the final theory examination.

Means and standard deviations were calculated for each set of marks and differences between means were checked for level of significance (see Tables 1 and 2). Comparing men with women, in 1968 there was only one significant difference, at the five per cent level, and that was for aesthetic value. In 1970 there were in addition significant differences for economic and political values also. Evans's graduates showed a much sharper difference between men and women. Assuming, as Evans did, that men and women should be treated as two different populations, changes over the two-and-a-half-year period were examined for each sex. In this case there were no significant differences except for a borderline decrease in religious value for women. It appears then that values are remarkably stable over a two-and-a-half-year period, while students are undergoing a period of training probably quite unlike anything they have experienced previously.

There is a possibility that the different courses, supervised by different tutors and with somewhat different aims and methods might have different effects on changes of values, small though they appear to be. To check this the students were divided into groups according to the age range of children for which the students were training. The 1970 results were extended by the inclusion of the specialist students who were not tested in

Table 1: Means and standard deviations—Study of Values and MTAI, 1968

VALUE	MEAN	S.D.	DIFFERENCE OF MEANS	LEVEL OF SIGNIFICANCE %
Theoretical	32·66* 32·02**	6·14 5·89	0·64	NS
Economic	32·25 30·11	6·30 5·45	2·14	NS
Aesthetic	24·47 28·65	7·96 7·13	4·18	5
Social	43·58 43·09	6·92 6·27	0·49	NS
Political	25·56 22·99	7·03 7·00	2·57	NS
Religious	21·58 22·95	14·04 10·95	1·37	NS
MTAI	32·94 39·57	21·68 24·70	6·63	1

*Upper Line – Men: N = 32
**Lower Line – Women: N = 112

Table 2: Means and standard deviations—Study of Values and MTAI, 1970

VALUE	MEAN	S.D.	DIFFERENCE OF MEANS	LEVEL OF SIGNIFICANCE %
Theoretical	33·81* 32·35*	5·35 5·93	1·46	NS
Economic	34·97 29·93	6·36 6·44	5·04	1
Aesthetic	25·70 29·90	7·22 7·33	4·20	1
Social	43·60 43·44	6·07 6·78	0·38	NS
Political	25·23 22·36	5·65 7·86	2·87	5
Religious	17·50 21·02	10·26 9·87	3·52	NS
MTAI	37·36 52·91	20·23 20·36	15·28	1

*Upper line—Men: N = 32
**Lower line—Women: N = 112

1968. Two analyses were carried out. In the first the change for each group over the two-and-a-half-year period was examined. In the second, groups were compared in pairs on the 1970 results and here the specialist groups were included as their courses differed considerably from those of the other groups. The men were excluded from these analyses as they did not extend over the whole range. Since the results showed that men were more flexible than women in their attitudes, their exclusion tends to enhance differences.

Again, considered in these groups, women students' values did not change over the two-and-a-half-year period (see Tables 3 and 4). When groups were compared on the basis of the 1970 results, only four significant differences were found: 'infant' students rated social value higher than did 'junior' and 'junior-secondary' students; 'infant-junior' groups rated social value higher than did 'junior-secondary' groups; and 'infant' students rated religious value higher than did 'junior-secondary' students.

The MTAI was adminstered at the same times as the Study of Values and the scores were taken as the algebraic sum of the positive and negative scores. The results were analysed in the same way as were those of the Study of Values. With nine significant differences out of a possible total of eighteen, this was obviously a field of greater change than was the Study of Values, especially as eight of these differences were significant at the one per cent level.

Women scored significantly higher (one per cent level) than men in both years and also increased their scores significantly over the two-and-a-half-year period, while the increase in the men's score over the same period was not significant.

Looking at the groups of women students, the largest increase in score was with the 'infant-junior' groups, while the 'junior' group's score was less on the second test. On arranging groups in order of scores the rank was 'infant-junior', 'infant', 'junior', 'junior-secondary' and specialist. The specialist groups scored significantly lower than all others, and the only other significant difference was between 'infant-junior' and 'junior'.

Summary of results

The body of students examined in this study showed little change of attitudes at the end of their course compared with the beginning. It

Table 3: Means and standard deviations—Study of Values and MTAI (Women only, arranged in 'Age Groups')

	Infants N = 53		Inf/Jun N = 26		Junior N = 26		Jun/Sec N = 9		Specialist N = 43	
Theoretical	29·92*	5·31	33·60	6·00	32·96	6·27	36·22	4·52		
	31·49**	4·16	33·69	5·04	33·62	8·06	31·37	7·79	31·76	4·52
Economic	29·58	5·15	29·33	6·12	31·50	4·44	32·61	5·49		
	30·01	6·26	30·00	7·07	30·94	5·60	29·89	7·29	31·78	5·14
Aesthetic	29·02	6·18	28·11	7·27	26·89	5·47	34·39	9·08		
	29·01	6·08	29·87	7·22	28·27	5·38	36·11	10·82	29·03	6·33
Social	43·66	5·95	42·92	6·38	42·73	6·80	41·44	5·04		
	45·92	5·29	44·27	5·37	41·23	8·15	39·44	8·13	43·48	6·04
Political	21·94	6·77	23·09	6·24	23·79	7·77	26·89	5·98		
	20·49	6·75	23·04	6·33	25·25	9·61	24·22	8·84	21·13	6·29
Religious	25·82	10·77	21·29	12·08	20·88	8·32	10·56	2·94		
	22·96	9·19	18·79	11·38	19·92	8·77	16·33	7·97	22·66	10·60
MTAI	45·62	24·95	26·69	24·66	37·72	18·53	46·67	21·06		
	53·21	21·21	60·08	18·94	37·15	16·15	48·22	23·28	24·23	25·44

*Upper line: 1968
**Lower line: 1970

would appear that they tended to be people who committed themselves early, and at the age of eighteen or before had already decided what they wanted to do as a career. They, both men and women, differed from the general public in having higher social values and lower religious values, of border-line significance with men, but highly significant with women (see Table 5). Women also showed a higher economic value—perhaps they were more career-conscious than most women. MTAI scores may be a measure of 'child-centredness', and both men and women students scored significantly higher than the general public. The women scored higher than the men and improved their scores very significantly over the two-and-a-half-year period. The scores and changes tended to be higher with students training for teaching younger children. In con-

Table 4: Level of significance of differences of means between 'Age Groups' (Women only) 1970—Study of Values and MTAI

Reference Group	Comparison Group	Social %	Religious %	MTAI %
Infant	Infant–Junior			
	Junior	5		
	Junior–Secondary	5	5	
	Specialist			1
Infant–Junior	Junior			
	Junior–Secondary	1		1
	Specialist			1
Junior	Junior–Secondary			
	Specialist			1
Junior–Secondary	Specialist			5

Table 5: Levels of significance of differences of means between students' scores (1968) and published norms for Study of Values and MTAI

VALUE	MEN %	WOMEN %
Theoretical		
Economic		1
Aesthetic		
Social	1	1
Political		
Religious	5*	1
MTAI	1	1

* Border-line only

versation and discussion it was fairly obvious that the students were 'on the side of the child'.

The results obtained indicated that if Study of Values were used in the selection of students for a teacher training course the results for all values should be considered as none seem susceptible to change over the period of the training course. On the other hand the results of the Minnesota Teacher Attitude Inventory did change significantly so this test would not make such a useful contribution to the selection process.

Examination of tests

On the basis of Spranger's definitions it would appear that teachers should be predominantly of the social type as so much in education depends on the daily interpersonal contacts between teacher and pupils. Concern for individuals should be tempered with a regard for truth, so theoretical values should be at least minimally present. People do not become educated if they ignore the truth, on the contrary, acknowledging the truth and all it implies is part of the educational process. Appreciation of form and harmony adds a warmth and a zest to life that no educated man should lack and the teacher who does not possess this appreciation cannot pass it on. At the basic, practical level the teacher must allow for economic values, but otherwise these need not be of any great importance. The religious type shows no qualities that mark him out as specially suitable for teaching, so religious values appear to be irrelevant. On the other hand political man could be downright harmful in the teaching situation if he is prepared to use people for his own ends. The dominating authoritarian teacher is now seen by many as a threat to the development of his pupils.

The meagre results thrown up by this analysis however led to a careful examination of the tests used. First of all Spranger described six types of men according to the values held. Like so many attempts at classifying, this was an attempt to simplify what is a complex situation. Allport (1963) pointed out that there is an obvious seventh type, the sensuous or hedonistic, which Spranger did not mention. It could be argued that any life style is based on a set of values so the number of types of men could be extended almost at will. The problem is to limit the number to a manageable size while maintaining useful types. Spranger's classification involves ambiguities too. For instance love of people may be founded on religious as well as social values; an appreciation of truth besides being based on theoretical values can also have aesthetic feelings. Spranger however did not insist on pure types and this is in accordance with common everyday experience. A more rigorous, philosophic discussion of sets of values might confirm Spranger's six types, or, on the other hand, might yield more or different types, and further work would be needed to relate these types to the activity of teaching.

Secondly, Study of Values has the inherent weakness that the testee is asked to report what he would do or think under hypothetical conditions, and so the validity of the test depends heavily on the sincerity of the testee. As Wilson (1972) pointed out, this is really a performance test and the same patent action may be performed for quite different reasons, and only a discussion with the testee can show what these reasons are. For instance, if a testee indicated that he would rather be a doctor than a minister of religion, say, it is really necessary to know the reasons for his choice before giving to it a score for either social or religious value. He may see a doctor as wielding a great deal of power and ranking high in the social order, which would make his motive political. He might be attracted by the high income that could be his as a doctor when the choice would be economic. He might be fascinated by the complex interplay of cause and effect in the working of the human body; if this fascination is an intellectual one then the value is theoretical; if the beauty of this complex and delicately balanced mechanism is the compelling attraction then the value of the choice is aesthetic; yet again he may see the human body as a wonderful result of creation with an accompanying religious feeling. So the one choice may be attributed to any one of the six values according to the underlying thoughts which the test does not elucidate, and the same sort of discussion could be applied to many of the tests.

Lastly, the test does not entirely match Spranger's definitions. Thus Spranger according to Allport (1963) and quoted in the test manual describes the theoretical type as follows:

> The dominant interest of the 'ideal' theoretical man is the discovery of truth. In the pursuit of this goal he characteristically takes a 'cognitive' attitude, one that looks for identities and differences, one that divests itself of judgments regarding the beauty or utility of objects, and seeks only to observe and reason. Since the interests of the theoretical man are empirical, critical and rational, he is necessarily an intellectualist, frequently a scientist or a philosopher. His chief aim in life is to order and to systematize his knowledge.

Checking against this description the test items elicit a preference for certain academic subjects, two of which are named in the description, a bias towards 'academic' work and research; a high opinion of intelligence and clear thinking; a liking for demanding work and making certain choices of careers and activities. The ordering and systemizing of knowledge, and the cultivating of empirical, critical and rational habits of thought, which appear to be just as essential qualities of theoretical man, may be implied in the preferences elicited by the test, but there should be a much more direct and obvious connection between the choices made and the values held. Similar arguments could be advanced about the other values.

The compilers of the Minnesota Teacher Attitude Inventory describe fully their concept of the right attitude for teachers to have towards children and the task of teaching. The final test for such an attitude is surely the effectiveness of the teaching performed, involving defining aims and assessing results. There is also the question as to whether different styles of teaching are better adapted to teaching different subjects. There was no obvious difference in the effectiveness of the teaching of the specialist students with older children and that of other students with younger children, though their scores on the Inventory were significantly different.

Since the Inventory contains a large number of items it tends to be reliable. It is also a performance test and so its validity is open to doubt. Evans (1958) found with a small experimental group of students that their scores changed significantly when they were asked, on re-testing, to give the responses that they thought most likely to please the examiner. So here again, the hidden reasons are more important than the overt responses.

Conclusion

This investigation has failed to show any useful application of Study of Values or Minnesota Teacher Attitude Inventory as a yardstick in the selection of students for teacher training. There appear to be some doubts about the validity of Study of Values in its present form, but this can be confirmed only by careful investigation into the motives of the testees for giving their individual responses. It would be desirable to know also whether responses in the Minnesota Teacher Attitude Inventory indicated sincerely held opinions or were given only to produce a favourable impression.

References

ALLPORT, G. W. (1963). *Pattern and Growth in Personality*. London: Holt, Rinehart & Winston, p. 296.

COOK, W. W., LEEDS, C. H. and CALLIS, R. (1951). *The Minnesota Teacher Attitude Inventory*. New York: Psychol. Corp.

EVANS, K. M. (1958). 'An examination of the Minnesota teacher attitude inventory', *Brit. J. Educ. Psych.* **28**, 255.

EVANS, K. M. (1967). 'Teacher training courses and students' personal qualities', *Educ. Res.* **10**, 1, 72-7.

RICHARDSON, S. (1965). 'The study of values', Test Manual. Slough: NFER.

WILSON, J. (1972). *Philosophy and Educational Research*. Slough: NFER.

Self-concepts Among Secondary School Pupils

Barbara Thompson*, formerly National Foundation for Educational Research

Summary

Approximately 500 out of a sample of over 2,000 first-year pupils in 20 secondary schools were identified by their teachers as being either particularly well-adjusted, maladjusted or as having appeared before a court. These pupils were asked to complete 'semantic differential' scales on which they rated for such concepts as 'myself', 'my mother thinks I am . . .', 'my father thinks I am . . .', 'my friends think I am . . .' and 'my teacher thinks I am . . .'. The investigation was repeated for most of the same pupils in the fourth year. There was little difference in the self-evaluations of the three groups, but interesting differences emerged between the groups as to how they believed others saw them.

The investigation was carried out in 1967 and 1970 as part of the NFER Constructive Education project, sponsored by the Home Office and the Department of Education and Science.

Introduction

It is generally accepted that the way an individual perceives himself influences the way he behaves in interaction with his physical and social environment. It may also be assumed that school children who are perceived by their teachers to be particularly well-adjusted or maladjusted, or who have a record of delinquent behaviour, will have different concepts of themselves, and will also vary in the ways they believe significant people in their lives perceive them.

In an extensive study of the self-image of over 5,000 New York adolescents, using a 10-item scale, Rosenberg (1965) found that those with low self-esteem were more likely to

(a) display a number of psychosomatic symptoms,
(b) feel it necessary to present a false front to the external world,
(c) be sensitive to criticism or attack,
(d) feel isolated and lonely,
(e) have difficulty in making friends,
(f) have little faith in people, and
(g) be shy and unwilling to initiate conversations.

Remarkably similar results were obtained by Coopersmith (1967) with a sample of ten-year-old boys. Thus low self-esteem would appear to be associated with the type of behaviour which is frequently assessed as maladjusted.

Reckless *et al.* (1956), found that pupils with a positive self-concept were unlikely to become delinquent, and introduced the notion of the self-concept as an insulator against delinquency. The results of one of their studies are summarized by Schwartz and Tangri (1965) as follows:

> The 'bad' boys see themselves as likely to get into trouble in the future; their mothers and teachers agree. The 'good' boys see themselves as unlikely to get into trouble; their mothers and teachers agree.

Their own study, using a 10-scale form of the 'Semantic Differential', confirmed that sixth grade boys designated as 'good' by their teachers had more positive self-concepts than those of their peers designated as 'bad'. When the self-concepts of the two groups of boys were correlated with their mothers', friends' and teachers' views of them, the 'good' boys' self-concept was only found to correlate significantly with the way their teachers saw them. The only significant correlation with the 'bad' boys' self-concept on the other hand, was with their mothers' view of them. This result can be interpreted in one of two ways:

(a) that the mothers of 'bad' boys fail to encourage independence, thus restricting the possibility that others can control the boys' behaviour by reflecting either negative or positive self-images;

---

*Currently Staff Tutor in Social Science, London Region.

(b) if the evaluations of the 'bad' boys made by their peers are not yet significant, and their teachers reject them (as they show by nominating them as 'bad') where but to their mothers can they turn for support?

Whichever interpretation is accepted the outcome remains the same; twelve-year-old 'bad' boys are over-close to their mothers and therefore likely to have difficulty in developing appropriate masculine identifications. In addition, Hurlock (1967), reviewing some of the work on delinquency, reports than an unrealistic self-concept is likely to be associated with delinquency since it increases the probability that the child

> will try to compensate for the feelings of inadequacy that come from falling short of an unrealistic self-image by behaviour that deviates from the socially accepted pattern (p. 485).

On the basis of the above findings, the following predictions can be made:

(a) *Well-adjusted pupils* will have positive and stable self-images;

(b) *Maladjusted pupils* will have less positive self-images, in particular they will feel themselves less 'good', 'reliable', 'successful', 'wise' and 'orderly' than well-adjusted pupils, and will be less certain of their sexual identification. As they are clearly having difficulties in school it will also be predicted that they will feel undervalued by their teachers.

(c) *Delinquent pupils* will also have less positive self-images than the well-adjusted pupils. However, unlike the other groups, their self-images are likely to be at variance with the images they believe others have of them. Where specific scales are concerned, it is predicted that delinquents will see themselves as less 'good' and 'reliable' and also have less clear sexual identification, but will see themselves as harder and stronger than other pupils.

(d) There will be a tendency for older pupils to have less positive self-concepts than younger ones, and this will be particularly so in the case of the opinions they believe their teachers have of them.

Method

The data to be analysed here were collected as part of a study of attitudes of first- and fourth-year pupils in 20 secondary schools carried out as part of the NFER Constructive Education Project, sponsored by the Home Office and the Department of Education and Science.

(a) *Subjects*

The sample, in 1967, comprised 1,310 boys and 1,073 girls in their first year of secondary school; and, in 1970, any of these pupils remaining in the same schools in their fourth year (1,170 boys and 949 girls).

Teachers were asked on each occasion to nominate pupils

(i) who stood out in their age groups as stable and well-adjusted;

(ii) 'whose "maladjustment" is such that you would undoubtedly refer them for psychological or psychiatric guidance or treatment if such a service would readily cope with the demands to be made on it . . . ';

(iii) who had appeared before the court at any time.

Table 1 shows the numbers of boys and girls in each group in their first and fourth years.[1]

(b) *The instrument*

Osgood *et al.* (1957), developed the 'Semantic Differential' as a measure of connotative meaning, i.e. that aspect of meaning concerned with the associations rather than the literal meaning of the word. Osgood's instrument attempts to register the connotations which a concept has for an individual on a number of bipolar scales such as 'bad' . . . 'good', 'unsuccessful' . . . 'successful'. The respondent has to place a mark in one of the intervals

Table 1: Numbers of pupils nominated as well-adjusted, maladjusted and delinquent

	Boys			*Girls*		
	Well-adjusted	Maladjusted	Delinquent	Well-adjusted	Maladjusted	Delinquent
1st year	164	64	21	147	44	6
4th year	165	40	70	174	49	14

---

[1] The data analysed in this paper refer to these pupils only.

on each scale. Thus 'peace' might be rated as very good but relatively unsuccessful. Such an instrument was considered to be appropriate to explore how pupils saw themselves and how they thought others saw them.

Questions have been raised as to the validity of self-ratings as a measure of self-concept (Coombes *et al.*, 1963). Whereas it is agreed that a certain amount of distortion will inevitably occur when a person is asked to present a view of himself to outsiders, it is considered that:

(i) an instrument such as the Semantic Differential, being less obvious in its intentions, will not be as subject to distortion as some of the instruments designed to measure self-concept more overtly;

(ii) a rating-scale of this type is the only way of getting comparable data on how a pupil views himself and how he believes others view him.

All pupils were therefore asked to complete a 16-scale form of the Semantic Differential for 12 concepts including: 'Myself', 'My mother thinks I am . . .', 'My father thinks I am . . .', 'My friends think I am . . .', 'My teacher thinks I am . . .'[*].

In a number of factor analyses of semantic differential data, carried out by Osgood and others (Osgood *et al.* 1957, Warr and Knapper, 1968), the scales have generally been shown to fall into three main clusters identified as Evaluation, Activity and Potency. These are taken to represent three major dimensions of meaning. As the instrument on this occasion was to be used as a measure of attitudes, the majority of the scales (eleven out of a total of sixteen) were selected to represent the evaluative dimension. This is the factor which accounts for the largest percentage of the variance and has been identified as the attitudinal component of meaning. The other five scales, three representing the potency dimension and two the activity dimension, were initially included in order to prevent pupils from identifying the purpose of the study too clearly.

Financial limitations precluded carrying out a factor analysis of the scales on the data collected in this study. Hence it has not been possible to check directly the actual relation of scales to dimensions. Previous research has shown (Osgood *et al.*, 1957) that the factors on which scales load vary according to the concept being rated. Evidence also exists that when judging particularly highly or lowly evaluated concepts all scores should tend to rotate towards a single dominant evaluative dimension. However, in the present analysis, the assumption has been made that the eleven scales selected to represent the evaluative dimension did indeed do so.

(c) *Measures derived*

Osgood suggests the use of two types of measures: (a) factor scores, and (b) 'D', or generalized distance, scores which indicate the overall distance between two profiles. A number of authors (Cronbach, 1958, Jackson, 1962, and Warr and Knapper, 1968) have criticized the use of the D statistic, as the formula can result in the same score being derived from a number of different relationships between profiles. Warr and Knapper state their preferences for analysing differences between groups in greater detail, and adopt a procedure of testing for significant differences between sets of responses to each scale. This procedure has been followed in the current study. In addition, a Total Evaluation score obtained by summing the ratings on each concept over the eleven evaluative scales has been used, plus a difference score consisting of the sum of the differences on each evaluative scale between the ratings of 'myself' and of the views of mother, father, friends and teacher.

(d) *Analysis*

The question as to whether it is appropriate to use parametric statistics for the analysis of semantic differential data has been discussed at length by Warr and Knapper (1968) who conclude that the three main requirements for the use of parametric statistics—similar variances, equal interval scales and normal distribution of responses—are normally met. All analyses quoted in this study therefore make use of the F and 't' statistics.

Results

(a) *Total Evaluation scores*

(i) *Myself*

As predicted, well-adjusted boys and girls have very positive self-concepts. First year well-adjusted pupils have a total evaluation score of

---

[*] Forms were completed for seven other concepts but these are not relevant to the current discussion.

61·35, giving an average position on the seven-point scale of 5·6; and fourth years, 59·88, with scale average 5·4.

First-year maladjusted pupils have significantly less positive self-concepts than their well-adjusted peers. However, by the time they reach their fourth years, the differences between the two groups no longer reach significance. Neither can their self-evaluation actually be considered low, as their average scale scores are still greater than 5·0, i.e. more than a point above the centre-point.

None of the delinquents' self-evaluations differ significantly from those of well-adjusted pupils.

There is a tendency, as predicted, for older pupils to have less positive self-concepts than younger ones, although the only comparison between first and fourth years to reach significance is that between all the well-adjusted pupils.

(ii) *Myself as others see me*

In their first year at secondary school, pupils who are nominated by teachers as maladjusted rate themselves as seen by their mothers, fathers, friends and teachers less favourably than well-adjusted pupils, and the differences are more marked for girls than for boys. By the time they reach their fourth year, the maladjusted pupils see their mothers' and friends' views of them as being favourable, as do those who are well-adjusted. However, their fathers' and teachers' views are still rated as significantly less favourable than those of the well-adjusted group. As shown in Table 1 only a very few first-year pupils (21 boys and six girls) were classified as delinquent. Having groups as small as this considerably reduces the probability of any of the comparisons leading to significant results. In fact, it is only when ratings for 'My teacher thinks I am' are compared for all delinquent and all well-adjusted pupils that a significant difference obtains, the delinquents perceiving their teachers' views of them as less favourable than the maladjusted pupils.

In the fourth year, the number of delinquent boys is 70, although there are still only 14 girls so classified. Delinquent boys rate their mothers', fathers' and teachers' views of them less favourably than do well-adjusted pupils, and these results also hold when ratings for boys and girls are combined.

Unlike the results for their ratings of themselves, there is a strong tendency for pupils' ratings of how others see them to decrease from their first to their fourth years. For well-adjusted pupils, it is their fathers and teachers who are now seen to value them less highly, while delinquent boys rate their teachers as well as both parents as having a less favourable view of them. No changes are found for maladjusted pupils.

When the change with age is looked at between the three groups of first- and fourth-year pupils of the same sex, it becomes clear that ways in which boys rate others as perceiving them deteriorates much more than those for girls. The older boys feel that their teacher and both parents value them less highly than when they were younger, whereas the change for the girls only reaches significance where their fathers' views are concerned.

(b) *'Myself'—other difference scores*

It was predicted that well-adjusted pupils would have stable and realistic self-concepts. If this were the case, it would be assumed that they would not see themselves as under- or over-valued by parents, friends or teachers. This prediction holds to some extent for the well-adjusted girls, although in their first year they feel undervalued by their teachers, and in the fourth year by both teachers and their mothers. The picture for the boys, however, suggests that they feel undervalued by everyone except their fathers. This exception, however, is an important one, as Rosenberg (1965) found that a close relation with father was a crucial factor in high self-esteem.

The only prediction relevant to this section concerning maladjusted pupils was that they would feel undervalued by their teachers, who nominated them as in need of treatment. This is supported for all maladjusted groups except first-year girls.

There is little support from the results for first-year delinquents for the prediction that their self-images would be unrealistic, that is that they would feel generally undervalued. Like the maladjusted pupils, both boys and girls feel undervalued by their teachers, and girls feel this also of their fathers. The prediction is, however, strongly upheld for fourth-year delinquent boys. They feel undervalued by their teachers and by both parents. The group of girls is again too small for significant differences to be likely. The result which comes nearest to significance for them shows a tendency for them to perceive their friends as overvaluing them.

(c) *Comparisons on individual scales*

(i) *Myself*

On the basis of previous work, predictions were made as to how particular groups of pupils would rate themselves on specific scales. As has already been stated, pupils did not vary greatly in their ratings of themselves. However, some significant scale differences were found when maladjusted and well-adjusted pupils were compared. The scale which best differentiated the two groups was 'unsuccessful–successful' (significant differences for first-year boys and boys and girls combined, as well as the combined fourth years). In addition, fourth-year girls rated themselves less 'good' than their well-adjusted peers, fourth-year boys less 'masculine' and the combined fourth years less 'reliable'. No significant differences were found on the other two scales for which predictions were made, 'foolish-wise' and 'muddled-orderly'.

Significant scale differences were even scarcer, when delinquent pupils' self-concepts were compared with well-adjusted ones. The only scales to differentiate these were weak-strong (first-year delinquent boys seeing themselves stronger than their well-adjusted peers), and 'soft-hard' where again delinquent boys in the first year and all fourth-year delinquents rated themselves as 'harder' than comparable groups of well-adjusted pupils.

(ii) *Myself as others see me*

It is reasonable to argue that the way an individual believes others see him has even more influence on his behaviour than the way he sees himself. Comparisons between the ratings of how well-adjusted and deviant pupils believe others see them give considerably more information on the characteristics on which these pupils feel they differ.

Maladjusted first-year pupils are best differentiated from the equivalent well-adjusted group by the scales 'dirty-clean' and 'unreliable-reliable', rating others' views of them as less favourable on both scales. They also feel they are seen as less good generally by friends and teachers, and less successful by fathers and teachers. Finally, boys and girls see themselves as rated by both parents further from the ideal sex-type than the well-adjusted pupils do.

The scale 'unreliable-reliable' is also the one which best differentiates older maladjusted from well-adjusted pupils. Only their mothers are not perceived as rating them as less reliable than the mothers of the well-adjusted pupils are perceived as rating their offspring. Fathers, friends and teachers are also perceived as finding them less successful, and the boys believe everyone rates them as less masculine. The general tendency of older maladjusted pupils to devalue themselves is also shown in the lower rating of them attributed to father and teacher on the scales 'boring-interesting' and 'bad-good'.

When responses of first-year delinquent and well-adjusted pupils are considered, the situation is relatively unclear. In fact, there is a tendency for first-year delinquent pupils to rate others' views of them as more favourable than well-adjusted pupils of the same age. This becomes highly significant in the scale 'weak-strong', where ratings of themselves as mother, father and friends see them are more positive than the same ratings by well-adjusted boys.

The scales which most differentiate the fourth-year delinquent pupils from their well-adjusted peers are 'soft-hard', 'cruel-kind' and 'feminine-masculine', delinquents feeling harder and more cruel and less clearly identified with their own sex. They are also seen as less clean, good and successful from at least three other viewpoints. Perhaps the most important information to be gleaned from the analyses of these pupils' ratings is how undervalued they feel by fathers and teachers. Differences from well-adjusted pupils reach significance for perceived views of fathers and teachers on twelve out of the sixteen scales.

Discussion

Secondary school pupils generally evaluate themselves very favourably, as indicated by the high group averages for the Total Evaluation Scores. The average scale score of the group with the least favourable self-rating (first-year maladjusted pupils) is still over 5·0, i.e. more than a point above the centre of the scale.

Contrary to prediction, however, self-ratings vary very little from group to group, only first-year maladjusted pupils rating themselves sufficiently differently from the well-adjusted group for the result to reach significance. It seems possible that even an instrument which conceals its intent as well as the Semantic Differential does, inhibits pupils from displaying negative feelings about themselves; and that only when they are asked to rate themselves as others see them are these inhibitions released sufficiently to display those differences which have led teachers to classify them as deviant. A further possible interpretation of this finding would be that in fact it is the perceived views of others rather than the degree of self-esteem itself which affects people's behaviour, but this we cannot even attempt to explore here.

Certainly the results from comparisons between the way various groups of pupils feel others see them are far more revealing. Not only do first-year maladjusted pupils, as would be expected from the earlier finding, see themselves rated less favourably than their well-adjusted peers by parents, friends and teachers, but the results of comparisons are also significant for the older maladjusted pupils when their rating of themselves as fathers and teachers see them are compared with those of the well-adjusted group. Differences between delinquent and well-adjusted pupils now reach significance where teachers' views are concerned in their first year, and for teachers and both parents in the fourth.

As the original classification of deviancy was made by teachers, the pupils are clearly realistic in perceiving them as undervaluing them. Both older groups also attribute less favourable assessments to their fathers, suggesting that their problems are not confined to school. The low evaluation attributed to their mothers by older delinquents contradicts the findings of Schwartz and Tangri (1965) that delinquent boys, at least, are closely identified with their mothers. However, the delinquent boys in this study are three years older and would appear to have moved towards closer peer group identification. Their friends are the only people whose evaluations of them do not differ significantly from the ways in which they evaluate themselves.

The finding that fourth-year delinquents feel undervalued by teachers and parents appears to confirm the prediction that their self-images are at variance with the views they perceive others to have of them. An alternative interpretation, however, would be that these pupils are peer- rather than adult-orientated. This interpretation seems better to explain the difference from the well-adjusted boys, who feel undervalued by mothers, teachers and friends, but not by their fathers. This finding is not only further evidence for the importance of paternal support for adjustment and self-esteem, but also suggests that identification with an adult figure may be important for perceived success in school. The fact that well-adjusted girls do not feel undervalued by fathers or friends would indicate that the positive role played by father is the crucial factor in good adjustment.

A number of scales differentiate the two older deviant groups from their well-adjusted peers. Both deviant groups perceive themselves to be seen as less successful, good, and masculine/feminine. However, the maladjusted pupils also feel themselves less reliable and interesting, scales which seem to indicate doubts about their acceptability as companions, while the other scales which differentiate the delinquents are part of the 'bad boy' stereotype—hard, cruel, dirty. The fact that they do not think they are perceived as masculine at the same time is a contradiction which gives rise to speculation as to the existence of some compensating mechanism.

Main points

1. Secondary school pupils assessed as well-adjusted, maladjusted and delinquent do not vary greatly when evaluating 'myself' on a form of the semantic differential.
2. Their ratings of themselves as others see them indicate that both the maladjusted groups and the older delinquents rate themselves as being perceived considerably less favourably than do the well-adjusted pupils.
3. Older delinquents appear to be peer rather than adult-orientated, in that only their friends are not seen to undervalue them.
4. The well-adjusted boys, on the other hand, feel that only their fathers do not undervalue them and this, combined with the fact that neither fathers nor friends are perceived as undervaluing the well-adjusted girls, underlines the importance of a positive adult identification.
5. The scales which appear to differentiate both deviant groups from the well-adjusted pupils are 'unsuccessful-successful', 'bad-good' and 'masculine-feminine'.
6. The scales on which maladjusted pupils also differ from well-adjusted pupils are 'boring-interesting' and 'unreliable-reliable', indicating doubts about personal value.

7. The scales differentiating delinquents on the other hand are 'soft-hard', 'cruel-kind' and 'dirty-clean', all part of the traditional tough 'bad boy' image.

References

COOMBES, A. W., SOPER, D. W. and COURSON, C. C. (1963). 'The measurement of self concept and self report'. *Educ. and Psych. Measurement*, **XXIII**, 3, 493–500.

COOPERSMITH, S. (1967). *The Antecedents of Self-Esteem*. San Francisco: W. H. Freeman & Co.

CRONBACH, L. J. (1958). 'Proposals leading to analytic treatment of social perception scores', in TAGUISI, R. and PETRULLE, L. (Eds.) *Person Perception and Interpersonal Behaviour*. Stanford University Press, Stanford.

HURLOCK, E. B. (1967). *Adolescent Development*. Maidenhead: McGraw Hill.

JACKSON, D. M. (1962). 'The measurement of perceived personality trait relationships', in WASHBURNE, N. F. (Ed.) *Decisions, Values and Groups*, Vol. 2, Pergamon Press, Oxford.

OSGOOD, C. E., SUCI, G. J. and TANNENBAUM, P. M. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

RECKLESS, W. C., DINITZ, S. and MURRAY, E. (1956). 'Self concept as an insulation against delinquency'. *Amer. Sociol. Rev.*, 21, 740–6.

ROSENBERG, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press.

SCHWARTZ, M. and TANGRI, S. S. (1965). 'A note on self concept as an insulator against delinquency', *Amer. Sociol. Rev.*, **30**, 6, 922–6.

WARR, P. B. and KNAPPER, C. (1968). *The Perception of People and Events*. New York: John Wiley & Sons.

Journal of Curriculum Studies **Collins**

In the November issue

One Project's Awkward Thinking about Objectives. *W. A. L. Blyth*

Curriculum Concepts and Conceptual Clarity. *D. K. Wheeler*

What can Teachers Contribute to Curriculum Development? *K. R. Gray*

The Politics of Political Education. *Guy Whitmarsh*

Curricula of Teachers' Colleges. *Philip H. Taylor and R. S. Adams*

Objectives in Geography Fieldwork. *David J. Boardman*

Nigerian Teachers' Reception of a New Syllabus. *P. A. I. Obanya*

It may be ordered direct from the Publishers

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada
APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

May and November
Annual Subscription £3.00 including postage

The Swing from Science

a Perspective from Hindsight

D. Duckworth, National Foundation for Educational Research, and
N. J. Entwistle, Department of Educational Research, University of Lancaster

Summary

The swing away from science was accepted as a disturbing reality in the report of the Dainton Committee. A recent research study has re-examined the inferences drawn from the statistical evidence of the swing and found that there has also been a similar swing away from foreign languages. It was subsequently found that there was a close relationship between the swing and the mean academic aptitude scores of pupils taking those subjects. In a cross-sectional survey in Lancashire the attitudes of pupils were shown to be rather more specific towards individual subjects than had previously been assumed. It was also shown that it was possible for pupils to be interested in a subject and yet to find it very difficult. Possible implications for science teaching at 'O' level emerged from these results.

Introduction

In the 1950s there was a marked increase in the proportion of pupils studying science in the sixth form. The reversal of this trend in the 1960's led to the setting up of the Dainton Committee which reported in 1968 (Council for Scientific Policy, 1968). The committee concluded that 'the proportion of "A" level pupils following science courses in the first year of the sixth form has declined steadily since 1962, while other groups (particularly the mixed group, taking both arts and science subjects) have grown. This is, in substance, the swing'.

The Dainton Committee sought explanations of the swing in terms of early specialization and teaching which was hampered 'by the weight of formal exercise and historical accretions'. But their main explanation for the swing hinged on changes in pupils' attitudes towards science. 'Clearly it would be desirable to establish whether there is real repugnance in the minds of many young people to the harder and more materialistic manifestations of science and technology.'

Immediately after the publication of the findings, science teachers had the opportunity to express doubts about the seriousness of the situation at a conference of the Association for Science Education. Newall (1969) reported that, to the extent that there was a swing, it could be attributed mainly to 'the complexity of the subject matter of science and the sheer hard work required in science courses'. Reporting the results of a Scottish follow-up study designed to establish factors relating to science choice, Butcher (1969) expressed similar doubts about the swing and its causes:

> In terms of numbers the swing has been relative, not absolute. Faculties of science have expanded but more slowly than other faculties. A similar trend at sixth-form level is clear. . . . So much for the common ground. When one turns to the causes for the swing, to extrapolation of the trend, to proposed remedies (or indeed the question of whether any remedy is required) no such agreement exists.

At about the same time as Butcher made these comments an inquiry was started at Lancaster into some problems associated with the swing away from science. To begin with there seemed to be some doubt about the precise interpretation of the statistical evidence presented by the Dainton Committee. A re-examination of the meaning of the official statistics was therefore carried out. Furthermore the conclusions of the Dainton Committee rested rather heavily on assumptions about pupils' attitudes. Thus a study of pupils' attitudes towards subjects in the second-year and fifth-year of secondary school was initiated.

The swing re-examined

Some of the confusion in discussions about the

Table 1: Passes in different 'A' level subjects as a percentage of total

		1963	1964	1965	1966	1967	1968	1969	1970	*Difference* 1970–1963
Main Science Subjects	Mathematics	16·8	16·9	16·4	15·6	15·1	14·3	14·0	14·3	−2·5
	Physics	12·8	12·7	11·8	10·9	10·5	9·8	9·4	9·3	−3·5
	Chemistry	10·0	9·8	8·6	8·3	7·7	7·3	7·2	7·1	−2·9
	Biology + Zoology	5·4	5·6	5·4	5·3	5·0	5·2	5·2	5·4	0·0
	Total Science	45·0	45·0	42·2	40·1	38·3	36·6	35·8	36·1	−8·9
Main Arts Subjects	English Literature	10·2	10·7	11·4	12·4	12·4	12·9	13·5	13·6	+3·4
	History	8·6	9·1	8·9	8·9	9·0	9·0	8·9	8·5	−0·1
	Latin	2·6	2·7	2·4	2·2	2·0	1·7	1·6	1·4	−1·2
	French	7·0	7·7	7·1	7·1	6·8	6·6	6·4	6·0	−1·0
	German	2·0	2·1	2·1	2·1	2·1	2·1	1·9	1·9	−0·1
	Total Arts	30·4	32·3	31·9	32·7	31·3	32·3	32·3	31·4	+1·0
Main Social Science Subjects	Economics	5·5	5·7	6·0	6·7	7·7	8·1	8·3	8·5	+3·0
	Geography	5·9	6·4	6·6	6·8	6·9	7·0	7·3	7·3	+1·4
	Total Social Science	11·4	12·1	12·6	13·5	14·6	15·1	15·6	15·8	+4·4
	Other subjects*	13·2	10·6	13·3	13·7	15·8	16·0	16·3	16·7	+3·5

Source: Statistics of Education, 1970, Volume 2, Table 32.

*General studies accounts for a substantial proportion of this category.

swing can be traced to the use of differing definitions. The swing can be measured in terms of the proportion of students entering the science faculties of universities, but this statistic depends on the number of places made available within universities and also ignores scientists studying in other types of institution. The most straightforward measure of the swing appears to be in terms of the proportion of passes obtained in each 'A' level subject compared with the total number of passes in all subjects. Statistics of Education (DES, 1970, Volume 2, Table 32) allows changes in the percentage passes in 11 GCE subjects to be identified between 1963 and 1970.

The first thing to note in Table 1 is that there appears to be a swing not only away from science, but also away from foreign languages. It is also apparent that the swing is not towards arts subjects in general (history, for example, barely holds its own), but towards English literature and the social science subjects in particular.

Apparently the evidence of distinct changes in the popularity of various school subjects is clear-cut. But is it? Changes can only legitimately be compared in identical situations. Was the sixth form of 1970 equivalent in all respects to the sixth-form of 1963 and, if not, were the differences likely in themselves to have exacerbated the swing? Certainly sixth-forms have changed in size, in composition and in curriculum; for instance, additional mixed arts-science combinations were becoming possible within an increasing number of school time-tables. Thus pupils whose commitment to science was less than total would in 1970 be able to choose one or two science subjects instead of the three they might have had to choose in 1963. Widening the choice of subject options available will inevitably bring about changes in the proportion of pupils choosing the 'traditional' subjects and subject combinations and it seems that some arts subjects as well as science subjects are affected by this expansion.

Another change in sixth forms is in the proportion of girls. Between 1965 and 1970 the number of boys in sixth forms increased by about eight per cent, while there were 30 per cent more girls (DES, Volume 1, Table 25). As McPherson (1970) pointed out, girls are less likely to choose science and hence an increase in the proportion of girls will inevitably lead to a decrease in the proportion of scientists. This explanation does not however, remove the swing. Among boys alone, there are substantial decreases in the proportion of pupils taking physics and chemistry, and similar decreases in the relative popularity of French and German (Duckworth, 1972).

The final change in the composition of sixth forms is in size. With more pupils staying on, many of these young people will be lower in ability than earlier generations of sixth-formers. As a School Council (1970) publication puts it, there is now 'a wider range of natural ability' in sixth forms. Changing social conditions certainly will have allowed some bright children from poor homes to continue at school longer than would have been possible seven years previously, but it would be surprising if the average level of ability had not been lowered by the rapid growth in numbers. If such a change is to affect the choice of particular subjects then pupils of lower ability must opt for different subjects from those chosen by the more able. This possibility does arise because there is a well-established mythology in schools about 'soft-options' at 'A' level, though only a little hard evidence to isolate particular subjects. Some evidence is provided in a study reported by Forrest, Smith and Brown (1970). Pupils taking Joint Matriculation Board 'A' level examinations completed a test of academic aptitude which gave scores on verbal and numerical ability. Looking at the mean scores of pupils taking the various subjects it is clear, for example, that pupils taking mathematics, physics or chemistry have higher verbal and numerical aptitude scores than those taking biology, economics or geography. Moreover, if we re-order the subjects shown in Table 1 in terms of the size of the percentage changes of the last column, there is a striking link ($\rho=0{\cdot}86$) between those changes and the total mean aptitude scores of pupils taking the subject (Table 2). Reservations must, of course, be made about the propriety of summing verbal and numerical scores equally to obtain a total aptitude score.

The percentage changes are not, strictly, measures of swing in individual subjects but they are closely related to them as is shown by a detailed analysis reported briefly elsewhere (Duckworth, 1974a). This analysis shows that the swing can be accurately described as a sharp decrease in the relative popularity of subjects for which higher aptitude scores are found. The 'new' sixth-formers of relatively lower aptitude are almost certainly not opting for physical science and language subjects.

This conclusion is supported by findings reported by both Phillips (1969) and Barnard and McCreath (1970). They show how early specialization in secondary schools restricts the options

Table 2: Comparisons between rank orders of subjects in terms of percentage changes (Table 1) and aptitude scores

RANK ORDER IN TERMS OF % DIFFERENCES 1970–1963	MEAN APTITUDE SCORES *Verbal*		*Mathematical*		STANDARDIZED APTITUDE TOTAL	RANK ORDER
	Raw.	Stand.*	Raw.	Stand.*		
Physics	36·8	33·7	32·1	46·9	80·6	1
Chemistry	35·5	29·9	28·9	41·6	70·5	4
Mathematics	34·6	27·2	32·8	48·0	75·2	3
Latin	42·3	50·2	19·8	26·6	76·7	2
French	37·0	34·3	17·3	22·5	56·8	6
German	40·2	43·9	18·5	24·5	68·4	5
History	34·7	27·5	15·9	20·2	47·7	9
Biology	33·3	23·3	21·3	29·1	52·4	7
Geography	30·3	14·3	17·4	22·7	37·0	11
Economics	31·3	17·3	20·7	28·1	45·4	10
English	35·0	28·4	15·7	19·9	48·3	8

*Raw scores standardized to a mean of 30 and standard deviation of 10.

available to pupils of lower ability. To be able to choose either science or foreign language 'A' levels a larger number of 'O' level subjects is normally required than for the choice of most other subjects. Again in most schools only the brighter potential sixth-formers are encouraged to carry seven or more 'O' levels and thus specialist study of science or foreign languages is often ruled out for the less able among them.

A similar but rather over-generalized conclusion about the relation between the swing and the changing composition of the sixth-form comes from reports of two local authority Education Committees.

'Heads felt that a major factor causing an apparent swing from science is that an increasing proportion of pupils is entering sixth forms and the increase comes from a relatively less able group who find the arts subjects easier of accomplishment' (Kent Education Committee, 1966).

'If less able youngsters, or those whose foundation work in mathematics/science is insecure, want to try "A" level, it is often possible with arts subjects but not with science subjects' (West Riding Education Committee, 1969).

While this re-examination of the statistical evidence considered by the Dainton Committee, together with more recent research evidence, throws doubt on the exact nature and meaning of the swing, the conclusions of the Committee about pupils' attitudes towards science may remain valid. The next stage of this investigation was designed to see whether the pupils themselves perceived science subjects as being uniformly more difficult, less interesting and having less social benefit than other subjects.

Attitudes of pupils

A repertory grid technique was used to measure four distinct attitudes towards nine school subjects. Full details of this method have been reported elsewhere (Duckworth, 1972; Duckworth and Entwistle, 1974). In essence the grid asked pupils to indicate whether a particular school subject could be allocated to one or other of opposed descriptive phrases: for example.

Tends to be difficult — Fairly easy

This item, not surprisingly, was part of the scale measuring the perceived easiness/difficulty of the various subjects. Three other scales were produced to cover the attitude dimensions of interest, freedom and social benefit. In a pilot study three attitudes only were investigated, but it was found by factor analysis and cluster analysis that 'freedom' should be defined separately and was linked in pupils' minds with the study of English. Key phrases defining this attitude were 'my own ideas can be used' and 'needs quite a lot of imagination'. Social benefit included items such as 'important for solving world problems' and 'knowledge useful in everyday life.'

This grid, together with other measures to be reported subsequently (Duckworth, 1974b), was given to a cross-sectional sample of pupils from the second and fifth forms of six Lancashire grammar schools. The sample consisted of 314 second-year pupils (145 boys and 169 girls) and 292 fifth-formers (134 boys and 158 girls). Mean scores for each sub-group were calculated for each of the four attitudes towards nine school subjects. As few pupils took all nine subjects, the mean scores were not uniformly derived from the same group. As pupils dropping, or not choosing, subjects are unlikely to have favourable attitudes towards them, this self-selection will tend to increase the measured attitude scores in less popular subjects such as Latin.

To simplify the presentation of the results rank-orders of the mean scores have been presented separately by age-group and sex in Table 3. The first point to note in these results is that any general statement about pupils' attitudes towards a group of subjects, such as the sciences, would be unwise. Pupils clearly hold rather different types of attitudes towards specific subjects within any such grouping, particularly in the second-year. For example, second-form boys see chemistry as interesting and easy, but physics as dull and difficult. Fifth-form girls see biology as interesting, free, relevant and not too difficult, whereas physics is again seen as being rather dull and very difficult. From Table 3 it is also clear that history, geography and English are uniformly perceived as being interesting and easy, although there are generally rather poor attitudes towards history in terms of freedom and social benefit. There are also consistently poor attitudes towards French and Latin and, even in the fifth form, these subjects come firmly at the bottom of the rank-order in terms of interest. It is interesting to note that the perceived difficulty of French decreases markedly between the second-year and fifth year, perhaps as the rudiments of the language have been mastered.

Table 3: Rank orders of school subjects for four attitude dimensions

	INTEREST		EASINESS		FREEDOM		SOCIAL BENEFIT	
	Year 2	*Year 5*	*Year 2*	*Year 5*	*Year 2*	*Year 5*	*Year 2*	*Year 5*
Boys	Chemistry	Geography	English	English	English	English	English	English
	History	Biology	Geography	Geography	Chemistry	Geography	Maths	Geography
	Geography	Chemistry	History	Biology	Geography	Chemistry	Geography	Biology
	Biology	History	Chemistry	History	Physics	Maths	Biology	Maths
	English	English	Biology	French	Biology	Biology	Physics	Physics
	Physics	Maths	Latin	Maths	History	Physics	Chemistry	Chemistry
	Maths	Physics	Maths	Latin	Maths	History	French	History
	Latin	French	Physics	Chemistry	French	French	History	French
	French	Latin	French	Physics	Latin	Latin	Latin	Latin
Girls	Biology	Biology	Geography	English	English	English	Geography	Biology
	Geography	Geography	English	Geography	Geography	Biology	Maths	English
	Chemistry	English	Biology	History	Chemistry	Geography	Biology	Geography
	History	History	History	French	Biology	Chemistry	English	Maths
	English	Chemistry	Maths	Biology	Maths	Maths	Physics	Chemistry
	Maths	Physics	Chemistry	Maths	History	Physics	Chemistry	French
	Physics	Maths	French	Latin	Physics	History	French	Physics
	French	French	Physics	Chemistry	French	French	History	History
	Latin	Latin	Latin	Physics	Latin	Latin	Latin	Latin

*This table is a modified version of a table already published in the *British Journal of Educational Psychology*, 1974, **44**, 79–80.

The fifth-form analyses are, however, obscured by most pupils by that time either having dropped subjects or at least having decided where their main interests lie. Additional analyses were therefore carried out in terms of the subjects which the pupils expected, or would have liked, to study in the sixth form. (Subsequent analyses based on actual choices altered the pattern of results only slightly). Table 4 summarizes the results on the two main attitude scales for fifth-formers choosing different groups of subjects. As only a few sex differences were detected, this analysis is reported for the age-group as a whole.

As expected the arts pupils and the science pupils favour their declared specialism, but the analysis produces more than this common-sense finding. Even among scientists, geography, English, and to a lesser extent history, are seen as easy and

Table 4: Rank orders of attitudes of all fifth-year pupils towards school subjects by subject choice

ARTS CHOICE		MIXED CHOICE		SCIENCE CHOICE	
Interest	*Easiness*	*Interest*	*Easiness*	*Interest*	*Easiness*
Geography	English	Geography	English	Biology	Geography
		Biology		Chemistry	English
History	Geography		Geography	**Physics**	Maths
English	History	History	History	Geography	Biology
				Maths	
		English	Maths	English	History
Biology	French	Maths	French	History	French
French	Biology	Chemistry	Biology		
Chemistry		**Physics**			Chemistry
	Latin				
Maths	Maths	French	Latin		
Physics	Chemistry		Chemistry	French	Latin
Latin	**Physics**	Latin	**Physics**	Latin	**Physics**

fairly interesting. Mathematics follows a predictable pattern, being thought dull and difficult by arts students and interesting and easy by scientists. However there is a complete contrast when the results for physics and chemistry are examined. Physics, in particular shows an unexpected pattern. Certainly arts students find it dull and difficult, but it is extraordinary to find that pupils who have opted for science subjects, many of whom will be taking physics at 'A' level, find it more difficult even than Latin, although they find the subject very interesting.

Discussion

Evidence from a variety of sources indicates that the 'swing from science' during the 1960's was not what it seemed at first sight. It was a swing away from subjects which were perceived to be difficult (the physical sciences and foreign languages) and towards easier and less restrictive subjects, English literature and the social sciences in particular. But the move away from the difficult subjects seems likely to have come mainly from less able pupils, pupils who would previously not enter the sixth form at all. In addition, the apparent swing away from certain science subjects has been artificially exaggerated by the increased proportion of girls among pupils taking 'A' level.

It seems arguable, therefore, that the sciences have not really 'lost' many recruits because of a change in pupils' attitudes towards science. Rather some have been attracted away by the welcome diversification of sixth-form courses while others, the 'new' sixth-formers, have never been serious candidates for science (or language) courses because they find these subjects particularly difficult. However, the striking finding that potential scientists find physics and chemistry difficult, even though they see them as interesting, must raise important questions for the designers of 'O' level syllabuses in these subjects.

There has been repeated pressure from some teachers to reduce the factual content of both 'O' and 'A' level science syllabuses. It may be that the problem goes deeper than this. With the rapid advance in the scientific disciplines are we now demanding, in addition to factual information, a degree of conceptual complexity which is unrealistic for fifth-formers to tackle? Perhaps more pupils would continue to study the physical sciences if the demands made at 'O' level were much reduced. But to make the method of treatment less taxing may, of course, only attract pupils who would be unable to cope with the conceptual complexity which would inevitably come later on. Nevertheless, some thought should surely be given to the question of the match between the intellectual maturity required in 'O' level pure science examinations and what can reasonably be expected from fifth-form pupils. There is also the disturbing possibility that the present tendency towards intellectually superior groups taking sciences and foreign languages, linked to a system which allocates approximately equal proportions of pupils to the various grades in each of the more popular 'A' level subjects, could be producing considerable variations in the standards of attainment required of those who are awarded particular grades in different subjects.

References

Barnard, G. E. and McCreath, M. D. (1970). 'Subject commitments and the demand for higher education', *J. Roy. Stat. Soc.*, **133**, 3, 358–408.

Butcher, H. J. (1969). 'An investigation of the "Swing from Science",' *Research in Education*, 1, 38–57.

Council for Scientific Policy (1968). *Inquiry into the Flow of Candidates in Science and Technology into Higher Education* (Dainton Report). Cmnd. 3541. London: HMSO.

Department of Education and Science (1970). *Statistics of Education*, Volume 2. London: HMSO.

Duckworth, D. (1972). 'The choice of science subjects by grammar school pupils'. Unpublished PhD thesis. University of Lancaster.

Duckworth, D. (1974a). 'The "swing" re-visited', *Educ. in Science*, 57, 32–3.

Duckworth, D. (1974b). 'Correlates of subject choice among grammar school pupils' (unpublished paper).

Duckworth, D. and Entwistle, N. J. (1974). 'Attitudes to school subjects: a repertory grid technique', *Brit. J. Educ. Psychol.*, **44**, 76–83.

Forrest, G. M., Smith, G. A. and Brown, M. H. (1970). *General Studies (Advanced) and Academic Aptitude*. Manchester: Joint Matriculation Board.

Kent Education Committee (1966). 'The Flow of Science Students to Higher Education' (unpublished report).

McPherson, A. (1970). '"Swing from Science" or "Retreat from Reason",' *Univ. Quart.*, **24**, 29–43.

Newall, R. (1969). 'Attitudes to science and scientists', *Sch. Sci. Rev.*, **51**, 174, 7–22.

Phillips, C. M. (1969). *Changes in Subject Choice at School and University*. London: London School of Economics and Political Science/Weidenfeld and Nicholson.

Schools Council (1970). *Proposals for the Curriculum and Examinations in the Sixth Form*. London: Schools Council Publication.

West Riding Education Committee (1969). 'West Riding Sixth Forms 1960/1 to 1968/9'. Unpublished Report.

Criteria for the Assessment of Pupils' Compositions

P. Berse, formerly National Foundation for Educational Research

Summary

An experiment was set up to discover whether ten- to eleven-year-old children would perform better on an imaginative writing task if given an aural stimulus consisting of an imaginary dialogue rather than the more commonly used written titles. The main energy of the investigation is focussed on the viability of different forms of assessing the quality of the resulting compositions. The most useful criteria from a large battery proved to be vocabulary, weighted index of modifying clauses, pre-verb length and sequence.

The experimental situation

The constructors of an English test for 10- to 11-year-old children decided to use a three-minute taped dialogue as a stimulus to the composition section of the test, rather than providing the conventional choice of one title from a list of four or five. The testing situation was used to develop and try out a variety of criteria for evaluating compositions. Using these criteria, the hypothesis that a taped stimulus would encourage 'better' writing than a title was examined.

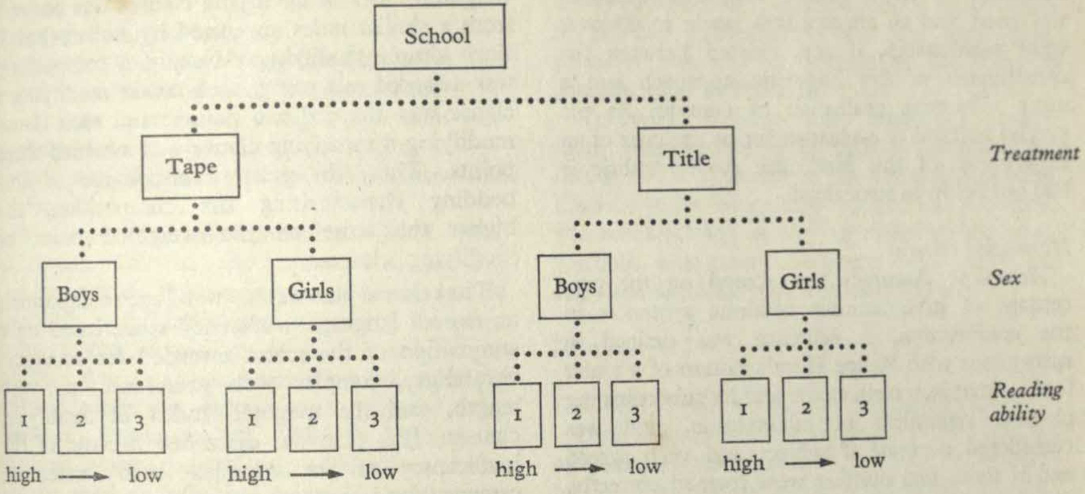
The tape was a conversation between a boy and his sister describing the strange events of the previous night during which all the furniture in the boy's bedroom had been mysteriously rearranged. The test instructions asked the pupils to listen to the story twice and think about how they might continue it in their own way. The title presented to the second half of the sample was 'Things that go bump in the night'. The test instructions requested that the children write something of their own using the title.

On the one hand, the experimental situation was designed to fit a practical problem, that of making an informed judgement about the format of a test. In order to make an assessment of the importance of the format, four possible sources of variance in the compositions were examined: school, sex, reading ability and experimental treatment (the way in which the composition section was presented). But the opportunity to compare the effect of two different types of presentation produced the more interesting and important problem of thinking out how compositions could be assessed. Various types of criteria were articulated and their interrelationship examined. Ancillary questions related to a range of features which might characterize the compositions of particular sub-groups of the population.

A sample of 447 children was drawn from the fourth-year junior group of eight schools representing a wide geographical distribution. A hierarchical design was used requiring 24 children from each school. One school did not contain sufficient children and was dropped at the analysis stage; also three of the schools were large enough to supply two replicates for the experiment. Thus the total design consisted of 10 'schools'. In each school there were two composition categories (Taped stimulus being T_1 and Title stimulus T_2), two sex categories and three reading ability categories.

The main effect of ability was controlled and interactions with ability were assessed. The reading ability categories were defined on the same absolute basis in each school, performance on NFER Reading Test BD. The observed distribution was divided into three ability bands with points of division at the 33rd and 67th percentiles. It was necessary to regard some pupils with scores near the cut-off points as representatives of the neighbouring ability group. There were seven cells for which it was not

The design



possible to make a plausible transfer and which were treated as a cell containing one missing value to be estimated from the other pupil in that cell. These cells did not then contribute to the between cells variance.

The total experimental population of 233 represented an equal number of boys and girls of each ability group in each school who received either the taped stimulus (T_1) or the title stimulus (T_2).

Random allocation to experimental groups was followed by instructions to the schools to administer the treatments simultaneously to all the children in the fourth year group, separating them for this purpose.

All children were given practice with taped stimuli during the week preceding the experiment to ensure that any novelty in the situation might be neutralized. Practice with a title stimulus was not regarded as necessary, as the children were already familiar with this type of presentation. The composition exercise took place at the end of a standard English comprehension test of the kind normally used in the examination procedure of the local education authority.

The analysis

There were two broad categories of criteria used to assess the compositions. The first set of criteria were normative; they were used to evaluate compositions along certain postulated dimensions of quality. The second set of criteria were descriptive; they were used purely to describe certain features of the compositions without implying any normative judgement. The criteria for assessing the compositions were based on counts of observed characteristics of the data; these were then divided into quintiles in accordance with the observed distributions in order to obtain equivalent measures for each for the sake of combination.

A four way analysis of variance was used to determine the relative contributions of treatment, sex, school, and reading ability to the variation of scores along normative criteria. A one way analysis of variance was used with each of the descriptive criteria in order to establish how different sub-groups of the population were responding to the experimental situation. Summary distributions for each variable were obtained. Correlation matrices were then formed for the constituents of the normative criteria and a factor analysis performed on these. The four way analysis of variance was then used to determine the sources of variation in the compositions in terms of school, sex, reading ability and treatment. Finally, the one way analysis of variance was used to determine characteristic features of compositions from different groups.

Criteria for evaluation

The task of comparing treatment effects was used to develop and explore ways of evaluating children's compositions, a job which is usually performed by the method known as 'impression marking'. A more precise linguistic approach was used and an attempt was made to discover what relationship, if any, existed between the constituents of this linguistic approach and a more subjective evaluation of content. As the precise method of evaluation forms the crux of an experiment of this kind, the coding rubric is laid out below in some detail.

Normative criteria

Accuracy: Accuracy was scored on the percentage of grammatically accurate sentences in the composition. A sentence was defined in accordance with Kellog Hunt's notion of a t unit (1965), that is, a main clause and its subordinating clauses, regardless of punctuation, and was considered accurate if subject and verb agreed and if tense and number were formed correctly. Spelling was not considered.

Vocabulary: A vocabulary score was determined by combining scores on four scales of vocabulary range: the number of different verbs, nouns, adjectives and adverbs.

Complexity

Complex verb preference: Complex verb preference was originally to be used in the manner suggested by Hess and Shipman (1972). The score was to be reflected as a ratio of complex verbs to the total number of verbs in the composition and this ratio then converted to a percentage. It was found, however, that the working definition of complex verb, which was any verb consisting of two or more words, made it impossible to distinguish between the continuous present and the pluperfect, thus rendering the category useless as a measure of linguistic sophistication. The category was discarded for the purpose of the analysis of variance, based on its complete isolation from other measures of complexity on the factor analysis.

Average pre-verb length: Average pre-verb length was computed by counting all the words preceding the main verb of each sentence in the composition and dividing this by the total number of sentences. Hess and Shipman suggest that this measure reflects a richness of modifying and qualifying language associated with complex and fluent speech. It was of interest to explore the possibility of applying this tool to written language.

Weighted index of modifying clauses: The weighted index of modifying clauses was adapted from a similar index suggested by Loban (1963). Each simple modifying clause in a composition was awarded one point; each clause modifying a clause was awarded two points; and each clause modifying a modifying clause was awarded three points. Thus, the greater the degree of embedding characterizing the composition, the higher the score on the weighted index of modifying clauses.

The original plan was to award each composition an overall language proficiency score based on a summation of the scores awarded for accuracy, vocabulary, complex verb preference, pre-verb length, and the weighted index of modifying clauses. This plan was discarded because of the inefficiency of the complex verb preference category, and separate scores were used in the analysis of variance for vocabulary, complexity (comprised of average pre-verb length and the weighted index of modifying clauses), and sequence (see below).

An effort was made to delineate a rubric for evaluating the content of the compositions for two reasons. First, it seemed useful to conceptualize some of the more intuitive ways in which a marker might respond to the substance of a composition. Second, it was of interest to explore the ways in which content and purely linguistic criteria might relate to one another.

Sequence: In attempting to isolate an aspect of content common to all the compositions, and yet meaningful enough to enable distinctions to be made between them, the notion of sequence presented itself as a relevant descriptive tool. 'Sequence' was defined as that aspect of a composition reflecting internal cohesion and organic development. By thinking in terms of a qualitative dimension of this kind it was possible to judge compositions of different sorts, for instance, stories with plots, poetic descriptions of some feeling or event, or dialogue representing dramatic presentations.

The usefulness of a criterion such as 'Sequence' depends as much on the sensitivity of the person evaluating the composition as it does on the inherent structure of the composition. This is an inevitable characteristic of any qualitative study of language and should be acknowledged as such.

The five points on the 'Sequence' scale were as follows:

1—sequence not attempted;
2—sequence attempted along one dimension;
3—sequence achieved along one dimension;
4—sequence attempted along more than one dimension;
5—sequence achieved along more than one dimension.

Possible sequential dimensions were plot, description, dialogue, and expressive association. The meaning of these terms needs to be clear in the minds of the individuals evaluating the compositions and shared to as great a degree as possible. Working definitions must be explicitly explored *within the context of the material being evaluated*. Examples of different degrees of sequence along the relevant dimensions must then be used as reference points for judging the sample of composition.

Descriptive criteria

Character relationships: This category did not apply to all the compositions, as some did not have characters, and was thus not used as a normative criteria. It proved, however, to be an interesting descriptive tool.

The scale was conceived along a dimension of relatedness. Characters were perceived as being presented as:

1—unrelated to one another;
2—related through ritualized conversation;
3—related through action;
4—related through feeling;
5—related through feeling and action.

Again, the issue of reader-injected meaning (or lack of it) was salient in using this rubric. After all, what is ritualized conversation and what conversation fluid and alive? These problems must be worked out within the particular context of the material being dealt with, taking into account such things as the age and experience of the children and the kind of situation in which they are being asked to write.

Type of response: Compositions were coded for the type of response they represented. These were defined as:

1—story (with achieved or attempted plot);
2—descriptive;
3—expressive or associative;
4—dialogue;
5—story plus dialogue.

The purpose in coding the compositions in this way was not normative, but meant solely to provide information about the form compositions assumed in relation to the kind of stimulus used to evoke them.

Remaining criteria: The remaining criteria again were used to gather information about the nature of the response to the stimulus situation. These criteria concerned whether or not there was scene setting at the beginning of the compositions, how many characters were used in the story and the length of the composition.

Results

Problem I: What are the sources of variance?

Differences attributable to treatment: The analysis of variance showed no significant differences between treatments in the three areas of composition evaluation. However, though vocabulary, complexity and sequence did not reflect treatment effects, there were certain criteria of composition description, rather than evaluation, which did.

Type of response: The tape evoked a greater number of story with dialogue responses than did the title among Ability Group 1 children at a very significant level for boys and a significant level for girls. There was a significant difference in the distribution of response types for the tape in Ability Groups 2 and 3. On inspection, it appeared that this was attributable to the high incidence of pure dialogue in these ability groups. The title evoked more Descriptive and Expressive-Associative responses across ability groups at a significant level for Ability Group 1 and a very significant level for Ability Groups 2 and 3.

Scene setting: Compositions were examined for the presence of an introductory statement which set the scene for the piece. Scene setting was significantly more characteristic of compositions written in response to the title than to the tape.

Character relations: There was a very significant difference in the distribution of types of character relations in the compositions. The title showed a higher incidence of Unrelated characters in Ability Groups 1 and 2 than did the tape in the same ability groups. The tape showed a higher incidence of characters related through Feeling and Action in Ability Groups 1 and 2 than did the title. The difference between treatments in the distribution of types of character relations was not significant in Ability Group 3.

Vocabulary means

Ability Group	Boys			Girls			Boys and Girls		
	T₁		T₂	T₁		T₂	T₁		T₂
1	12·56	←	12·65	15·00	←	15·84	13·81	←	14·21
2	11·30	→	10·70	13·55	←	13·90	12·42	→	12·30
3	9·05	→	9·00	11·37	→	8·83	10·18	→	8·92

Complexity means

Ability Group	Boys			Girls			Boys and Girls		
	T₁		T₂	T₁		T₂	T₁		T₂
1	5·83	←	6·30	5·63	←	6·58	5·73	←	6·44
2	6·10	→	5·95	5·85	←	6·20	5·97	←	6·07
3	5·20	→	4·80	5·05	←	5·39	5·13	→	5·08

† The arrow points in the direction of the lesser mean score.

Weighted index of modifying clauses: The weighted index of modifying clauses was one constituent of the complexity score. Although the complexity scores did not reflect treatment differences, the clause scores did, though only with one group. There was a highly significant difference in Treatment responses for Ability Group 2. In this ability group, the tape scores on the clause index were lower than the title scores.

Although the analysis of variance did not show treatment differences, an examination of the means for vocabulary and complexity may indicate some possible aspects of treatment effects.

It would appear that the means for boys on vocabulary and complexity show a gradual change from the tape being less than the title in Ability Group 1 to the tape being more than the title in Ability Group 3. The means for girls are less in the tape situation than in the title situation, with the exception of Ability Group 3 girls in the tape situation.

Differences attributable to sex: The analysis of variance revealed a very significant difference between sexes on vocabulary and sequence scores. No sex difference was evident on complexity scores. Differences between boys and girls were reflected through the use of descriptive criteria as well as along normative dimensions.

Length: Girls wrote longer compositions than boys, at a very significant level of difference.

Scene setting: Whether or not children set the scene for their compositions tended to vary in a different way for boys than for girls according to ability. Whereas Ability Group 1 boys set the scene more often than boys in Groups 2 and 3, girls' scene setting did not vary with ability. Girls in Ability Group 2 set the scene at a different rate from boys in Group 2, that is, they did so more often; however, the differences between boys and girls were not significant in Groups 1 and 3.

Stimulus utilization: The tape presented four stimulus characters to the children. Their compositions were examined to see how many of these characters were utilized. Though differences between boys and girls were not significant in Ability Groups 1 and 2, Group 3 girls more often used all four stimulus characters than did Group 3 boys. This difference was highly significant.

Character relations: Ability Group 2 girls responding to the tape conveyed character relationships through feeling and action significantly more often than boys in the same group. The differences were not significant for other groups.

Differences attributable to ability groups: The analysis of variance demonstrated highly significant differences between ability groups in vocabulary and in sequence. It did not demonstrate significant differences between ability groups on complexity. Though these differences were there, they were not large enough to reach significance level.

Length: Higher ability children wrote longer compositions.

Response: Higher ability children tended to write story with dialogue, while low ability children tended to produce straight story responses.

Stimulus utilization: Given four characters in the tape, the higher ability children more frequently used all of them than the low ability children.

Character relations: Higher ability children more often wrote compositions with characters related through feeling and action. This was very significant for boys and highly significant for girls.

Pre-verb length: Taken on its own, the distributions for the three ability groups were very significantly different according to the Chi-square test. Another test procedure (Kolmogorov-Smirnov) was used to test whether the results from one ability group had higher scores than each of the others. The only significant comparison was between Group 1 and Group 3, showing that Ability Group 1 produced compositions with a longer average pre-verb length than Ability Group 3.

Weighted index of modifying clauses: Compositions from Ability Group 1 received the highest clause scores, and those from Ability Group 3 the lowest. This was a very significant difference.

Differences attributable to school: The analysis of variance showed a highly significant difference between schools on vocabulary and a very significant difference on complexity. The experimental design removed this source of variance before analysing the effects of treatment, sex and ability.

Interaction effects: The analysis of variance revealed no significant interaction effects between treatment, sex and ability.

Problem II: What are the relationships between criteria?

Correlations: There were originally five criteria contributing to the linguistic analysis of the compositions. The highest correlation showed by accuracy with any other of the linguistic criteria was ·18. This was not regarded as meaningful.

All contributing elements to the vocabulary score correlated highly with each other (each correlation reaching at least ·52) and with both the weighted index of modifying clauses (·48) and with sequence (·58).

The maximum correlation for complex verb preference with any other individual criteria was ·06. This was not regarded as meaningful.

Pre-verb length correlated with the index of modifying clauses (·39).

The weighted index of modifying clauses correlated with vocabulary (·49), sequence (·33) and pre-verb length (·39).

Factor analysis: The factor analysis confirmed the information yielded by the correlations.

Three factors emerged. Factor 1 carried vocabulary plus the clause index and sequence. Factor 2 carried pre-verb length and the clause index. Factor 3 showed complex-verb preference on its own.

Examination of means contributing to the analysis of variance: Insight into the relationships between criterion scores was sought by examining the mean scores on individual criteria for linear trend.

Length: Complexity appeared to increase as the length of the composition increased. This was highly significant. With less certainty, it appeared that content development scores also increased with the length of the composition, as did vocabulary scores.

Response: The set of mean vocabulary scores varied very significantly according to the type of response. The highest mean vocabulary score occurred among the story with dialogue compositions. The lowest mean score came from dialogue alone. Dialogue also had the lowest mean complexity score, showing a highly significant difference from the remaining response types.

Character relations: As character relationships became more complex, that is, moved from unrelated to related through feeling and action, the vocabulary mean increased. This was highly significant. There was, however, little difference in the mean vocabulary scores for responses falling into the categories of unrelated and related through ritualized conversation.

Interpretation

Problem I: The sources of variance

Differences attributable to treatment: It must be remembered that no conclusions about tapes in general can be drawn from this experiment. What has been learned about the children's responses has been learned about this particular tape in a testing situation. Many of the inferences drawn from the data are related to the dialogue form of the tape itself.

Type of response, scene setting and character relations: It can be seen from examining the results in these three categories that the nature and form of the tape impresses itself upon the children's compositions, but that it does so in different ways for children of different reading ability. Ability Group 1 children appear to have taken fuller advantage of the taped stimulus by integrating to a greater degree two types of response in their compositions, that is, story plus dialogue. Ability Groups 2 and 3 appear also to have been affected by the presence of dialogue in the tape, but to have used it in a more imitative rather than innovative way.

Treatment 2, presenting a title to the children, was a less structured and more open invitation to write. Thus we see that there is more scene setting in this situation across ability groups and a greater number of expressive-associative responses. Where there were poetic creations, such as one girl's piece about a dream of talking crockery, they tended to be written in response to the title.

The tape, however, seemed to offer the children in Ability Groups 1 and 2 an opportunity to express a wider and richer range of character interaction than did the title. Given interesting voices and the beginnings of an exciting plot, these children were able to convey relationships with some substance. Ability Group 3 children hearing the tape did not convey character relationships in a manner distinct from that of their peers receiving the title. This may be interesting for teachers who are especially concerned with stimulating the less able child. Often the assumption is made that the more elaborate the stimulus, the richer the response. What may need to be considered is not only the extent of exposure to a stimulus but also the ways in which the teacher can act as mediator between the child and the experience, giving him maximum opportunity to absorb and integrate what is new.

Weighted index of modifying clauses: Noun clauses were not used in scoring the compositions on the clause index. They were omitted because it was expected that the children hearing dialogue on the tape would be predisposed to write reported speech in their compositions (John said that . . .) and that this would distort the index by reflecting a preponderance of noun clauses in Treatment 1 compositions which were due more to form-imitation than to spontaneous sophistication in language use. As a result of compensating for distortion in this way, the opposite picture emerged. Compositions written in response to the title, which were largely narrative, had higher clause scores than those written in response to the tape, which were largely characterized dialogue and reported speech.

Vocabulary and complexity means: Although what has been said about treatment effect in terms of type of response and character relations would seem to indicate that the better readers were getting more out of a taped stimulus, there is some slight evidence to suggest that in terms of vocabulary and complexity, the lower reading groups seemed to benefit. Boys in Groups 2 and 3 have higher mean vocabulary and complexity scores with the tape than with the title. Ability Group 1 boys do better with the title. With girls the situation is different, the tape providing an advantage only for Group 3 girls on vocabulary. The differences were not large enough to be statistically significant and are mentioned here only as a possible aspect of treatment effect. It is clear, however, that boys and girls are responding to the situation in different ways.

Differences attributable to sex: With regard to the comparison of means between Treatment 1 and Treatment 2 discussed above, it is possible that the content of the tape is more inspiring for boys than for girls. An attempt was made to avoid this by choosing a taped story with both male and female figures actively participating.

However, it is the boy who is the centre of the action, the girl taking a lively but subordinate role in the plot. Perhaps it is the boys' identification with the 'doer' in the story which makes them write with richer vocabulary and greater complexity in response to the tape as opposed to the title. This would not explain, however, either the fact that Ability Group 1 boys do better with the title or that the means for sequence do not show the same trend.

It is generally accepted that girls of this age are more linguistically proficient than boys. This proficiency may be particularly reflected in response to an unstructured situation.

Length, scene setting, stimulus utilization, character relations: The results in these categories of evaluation support the evidence of the analysis of variance, namely that girls do better than boys on vocabulary and sequence. They appear to write more and with greater elaboration in providing a setting and using and developing characters.

Differences attributable to ability: Results confirm the viability of the criteria used to evaluate the compositions in that one would expect the better readers to achieve higher scores on categories reflecting 'goodness' in compositions. However, the analysis of variance did just fail to show a significant difference between ability groups on complexity. On further investigation, this appears to be attributable to the high proportion of Ability Group 2 compositions with a high pre-verb length. The investigators suspect that pre-verb length may not only reflect linguistic complexity, but may also indicate a certain kind of wandering and repetition, characteristic of writing which is lacking in mature conciseness, but structured enough to be fluent and coherent. If this were true, the effect would create some contradictory movement among the complexity scores, thus making it difficult to see clearly the role of ability in determining the scores.

Problem II: Relationships between criteria

Accuracy, as measured in this experiment, had very little to do with the other indices of 'goodness' in compositions. It appeared as a kind of general non-directional influence in the factor analysis, bearing no meaningful relationship either to linguistic or content criteria.

The correlation matrix and the factor analysis revealed that one of the linguistic criteria, complex verb preference, was unrelated to every other dimension of evaluation, and so it was discarded as a constituent criterion.

A complex verb had originally been defined as any verb composed of more than one word, which forced the inclusion of the continuous present as a complex verb. This nullified any distinctions the count might have made.

The weighted index of modifying clauses appeared both on Factor 1, vocabulary and sequence, and on Factor 2, pre-verb length. This appears to indicate that the clause index is in some sense, two dimensional. It would be interesting to see if this two dimensional aspect were maintained when the index included noun clauses.

Main findings

1. There was no statistically significant effect of treatment on compositions.
2. Girls appear to be more proficient than boys on a majority of criteria, though they appear to benefit less from the tape stimulus than the boys.
3. Higher ability children score more highly on most of the criteria.
4. There is some interaction, between treatment and ability for boys. Ability Group 2 and 3 boys appear to benefit more from the tape than do Ability Group 1 boys.
5. The correlation matrix and the factor analysis indicate that the most useful criteria for composition evaluation used in this experiment were vocabulary, weighted index of modifying clauses, pre-verb length, and sequence.

References

HESS, R. D. and SHIPMAN, V. (1972). 'Early experience and the socialization of cognitive modes in children', in CASHDAN, A. and GRUGEON, E. (Eds.) *Language in Education: A Source Book*. Open University Press, Routledge & Kegan Paul.

HUNT, K. (1965). *Grammatical Structure Written at Three Grade Levels*, Research Report No. 3. Champaign, Illinois: National Council of Teachers of English.

LOBAN, W. D. (1963). *The Language of Elementary School Children*. Research Report No. 1. Champaign, Illinois: National Council of Teachers of English.

A Review of Three Journals for 1968-70

T. Derrick, School of Education, University of Bradford

Summary

As in the previous investigation (Derrick, 1973), there were the same considerable differences in the distribution of the articles amongst the age groups of the persons studied, the sizes of the samples, the methodologies of the designs, and the number contributed by individual British universities. Education was of minor interest in the sociology journal and British psychologists were quite uninterested in contributing to the international journal.

Introduction

Following the publication of the article 'Educational research: a three-year review, 1968-70' (Derrick, 1973) similar reviews of three other journals for these years were carried out, as the second stage of the project. It was considered that the particular journals selected would add to the picture of work done in British Education in the areas of psychology and sociology, and internationally.

The journals chosen were *The British Journal of Educational Psychology*, *The British Journal of Sociology*, and the *International Review of Applied Psychology*. The British Journal of Educational Psychology is an independent journal of a limited company in which the British Psychological Society and the Association of Teachers in Colleges and Departments of Education hold equal shares with an editorial body drawn from the university section of these bodies. The British Journal of Sociology is published for the London School of Economics, and the International Review of Applied Psychology is the official journal of the International Association of Applied Psychology. (This association records 178 British members (6 per cent) from an international membership of 2,800 from 90 countries.)

The methodology and definitions adopted were identical with the *Educational Research* review (Derrick, 1973).

Acknowledgements

Publication of this article was conditional upon its being in a factual version acceptable to the editors of the journals reviewed. Draft copies were sent to the three editors concerned and their provision or correction of some factual information was very helpful in the interests of accuracy. The interpretation of the data and their discussion are, of course, the sole responsibility of the author.

Results

The costs of the journals

Table 1: Cost per 100 pages in 1968-70

	1968	*1969*	*1970*	*Increase*
British Journal of Educational Psychology	16/7 (83p)	16/7 (83p)	20/7 (£1·03)	23% (over 10 years)
British Journal of Sociology	15/6 (77½p)	15/6 (77½p)	15/6 (77½p)	—
International Review of Applied Psychology (including membership of the association)	20–25/- (£2–£2·25)	20–25/- (£2–£2·25)	20–25/- (£1–£1·25)	—

The subject matter

Table 2 lists the number of articles which could be described as being of some interest to educationists in these journals in the three year period, 1968-70.

Table 2: The number of 'Education' articles in each of the journals, 1968-70

	Number assessed as 'Educational'	*The total number of articles in each journal*
British Journal of Educational Psychology	81	81
British Journal of Sociology	6	100
International Review of Applied Psychology	18	52

The articles

Table 3: The types of methodology used

	Comparative Survey	*Experimental*	*Theoretical*
British Journal of Educational Psychology	68	15	6
British Journal of Sociology	1	0	5
International Review of Applied Psychology	9	4	5

Table 4: The ages of the people studied

	1–7 yrs.	*8–11 yrs.*	*12–18 yrs.*	*Adults*
British Journal of Educational Psychology	14	31	38	14
British Journal of Sociology	2	2	6	2
International Review of Applied Psychology	0	6	7	8

Table 5: The numbers of people studied

	20–30	*40–80*	*100–250*	*300–600*	*1000+*
British Journal of Educational Psychology	3	13	25	20	17
British Journal of Sociology			1		
International Review of Applied Psychology	1	3	4	3	1

British Journal of Educational Psychology

Methodology

The comparative survey type of article predominated, four-fifths being in this category; theoretical articles were fewest in number, and were contributed by universities; the experimental articles were (proportionately) more popular with non-education sources than with British universities generally (see Table 3).

The Editor in his 'Notes to Contributors' encourages replication of experiments and surveys but evidence of this seems rare. The following is an exception: 'In retrospect this study suffered from two major defects. The first was lack of replication. . . .' (Connaught and Skurnik, 1969, p. 232).

Age

Amongst the subjects investigated, the youngest and the oldest age groups were under-represented; the youngest especially so in British universities compared with foreign universities, and the adult subjects vice-versa. Non-education sources seemed less interested in older subjects generally (see Table 4).

Experimental articles seemed primarily concerned with eight- to 16-year-old subjects, but the total number of articles in this category was only 14.

Sampling

There were few articles with samples of less than 40 subjects, or more than 7,000. One article concerned one subject. Comparative surveys usually sampled larger numbers than did experimental designs. Amongst the age groups the youngest and the oldest were under-represented in the larger samples. A characteristic which was often found in the articles of the most frequent contributors to the *British Journal of Educational Psychology* was the use of large samples (see Table 5).

British Journal of Sociology

Most of the articles were of the theoretical/historical type and they ranged over the whole scale of subjects' ages.

International Review of Applied Psychology

Half the articles were of the comparative survey type, and the rest experimental or theoretical/historical. Children younger than seven were not studied. The samples ranged widely but, as before, experimental designs had smaller samples than comparative surveys.

The authors

British Journal of Sociology

There were six authors, three of whom came from British universities (Aberdeen, Bath and Nottingham), two came from one US university, and one from a London college.

International Review of Applied Psychology

Universities in Europe and North America contributed the majority of the 18 education articles; France contributing seven and Great Britain none. Most of the above articles were by single authors.

British Journal of Educational Psychology
Collaborators

Half the number of articles had been written by single authors. British universities provided most of the articles, and schools the fewest. No British university had collaborated in joint articles with foreign universities; and only in two articles did a British university writer collaborate with one from another British university.

Number of contributions

Most of the authors had contributed one article only, but one British university author, and one Agency author, each appeared four times. Both these writers were notable for having used samples of more than 3,000 subjects.

Employing institutions

A comparison of British universities was made according to the numbers of articles they had contributed. The University of London provided eleven, Manchester nine, and Birmingham seven articles. Aberdeen, Bradford, and Keele supplied three each. Cambridge, Lancaster, Leeds, Nottingham, Southampton and Swansea, two each. There was one each from Belfast, Bristol, Edinburgh, Exeter, Leicester, Oxford, Sheffield and York.

Discussion

The findings of imbalance in research methodologies, the ages of subjects, the numbers of articles from different universities, and the lack of collaboration between universities, are very similar to those found in the *Educational Research* review (Derrick, 1973). On the question of the scarcity of experimental articles which was commented upon there, Eva Holmes (1972) writes similarly:

> 'Reading this collection from beginning to end leaves one feeling rather depressed about the nature of educational research. Everyone is working very hard, ploughing his own furrow and providing a great deal of useful information, but much of it is news, rather than science: . . . This umbrella usage (of the term 'research') tends to obscure the fact that controlled evaluative research . . . is rare.'

as do Trent and Cohen (1973) concerning US research on higher education teaching:

> 'Experimental studies in which significant variables have been systematically manipulated are exceedingly rare.'

As in the *Educational Research* review, the very young and the adults did not receive their fair share of attention. The particularly poor record of the *International Review of Applied Psychology* concerning the very young was especially noticeable.

The value to educationists of both the *International Review of Applied Psychology* and the *British Journal of Sociology* is possibly low since they published few education articles.

In the 1973 article it was written:

'That a minority, instead of a majority, of

British universities were represented by articles is a state of affairs which may well be disadvantageous, both to the journal and to universities as a sector of public education. This survey is one in a series on British education journals and it may, or may not, be found from other journals that there are 'horses for courses', i.e. special relationships, academically desirable, between certain journals and certain sources of contributions' (Derrick, 1973).

This present survey found no support for the 'horses-for-courses' hypothesis. Only a score of British universities had contributed to *Educational Research* and it was the same universities who contributed to the *British Journal of Educational Psychology*. It is difficult to write constructively about the meagre contribution of British universities to education in the *British Journal of Sociology*.

One could write with surprise that half of Britain's universities had not contributed to articles in the *British Journal of Educational Psychology*, or in *Educational Research*. Or of the complete absence of British universities in the international journal and their near absence in education in the sociology one. But what is claimed to be surprising is not these statements in themselves, but that no one seems to have made them before, or if they did, that they made so little impression in so doing. The Dainton Working Group on the Rothschild customer/contractor proposals (1971) claimed that 'a large contribution to research and to the education of future scientists is made by academic staff in science-based university departments' (p. 4), and this view was supported by the Social Science Research Council (SSRC, 1972). Nevertheless, the retiring Chairman of the SSRC Educational Research Board complained in the same letter that 'most academic departments in British Universities appear to be largely uninterested in any educational problems whatsoever other than those which affect universities' (ibid. p. 13) and there is no evidence in later issues of the SSRC Newsletters that anyone at all in the communities of scholars felt maligned by this stricture.

It is just perceptible that these views of the Dainton Committee and of the SSRC are being challenged, albeit anonymously, in the Department of Education and Science's *Trends in Education* (1972).

The third stage of this survey will be concerned with the relationship of educational research and educational practitioners.

References

CONNAUGHT, I. M. and SKURNIK, L. S. (1969). 'The comparative effectiveness of several short-cut item analysis procedures', *Brit. J. Educ. Psych.*, **39**, 3, 225–32.

DAINTON, F. (1971). 'A report of a C.S.P. Working group under the chairmanship of Sir F. Dainton', in *A Framework for Government Research and Development*. London: HMSO.

DERRICK, T. (1973). '*Educational Research:* a three-year review (1968–70)', *Educational Research*, **15**, 2, 94–6.

HOLMES, E. (1972). 'Review of *Educational Research in Britain—2*', *Brit. J. Educ. Psych.*, **42**, 1, 92.

SSRC (1972). Council's submission to government on *A Framework for Government Research and Development* (Cmnd. 4814), SSRC Newsletter 14, March, 2–6.

TRENDS IN EDUCATION (1972). 'Educational research and development', April, 2–3, DES and HMSO.

TRENT, J. W. and COHEN, A. M. (1973). 'Research on teaching in higher education', in TRAVERS, R. M. W. (Ed.) *Second Handbook of Research on Teaching*, 1052. New York: Rand McNally and Co.

Research Note

'Interpreting Research: Sampling'

A Reply

Gerald A. Cortis, School of Education, University of Birmingham

Choppin (1974) in his recent paper makes a valid point about the tendency for readers to accept any 'scientific' conclusion that appears in print. The same strictures of course apply to his own comments. I shall deal only with my paper (Cortis, 1973) since I am sure that Douglas (1973) and Quigley (1973) whose work he also analyses are more than capable of making their own case if they choose.

It is not clear from his comments whether Choppin is complaining that my work is inadequately reported as to sampling or whether the sampling itself is inadequate. The two points are quite different. At the outset he charitably suggests space as a reason for my supplying limited sample details and is right in this but immediately launches into a mixture of opinion and supposition somewhat removed from his own exhortations 'to approach . . . critically and with caution' (p. 218).

First he suggests that the sample 'is not as inadequate as was suggested in the Press' (p. 219). Only one report, as far as I know (not 'the Press' which suggests a number), said it was inadequate for reasons far removed from research ones and some reports noted how carefully the terms of reference were defined.

Second, the original group (respondents plus non-respondents) that he feels should be representative comprised 955 teachers from one of the larger ATO's in the country representing 6·90 per cent of the total output (13,854) of three-year and one-year graduate trained teachers in that particular year (1965) from colleges of education (DES, 1968, pp. 12-13).

Third, he suggests I have ignored sex differences and distributions amongst teachers. Nationally the number of women teachers in primary and secondary schools in 1967 was 172,876 whilst men teachers totalled 126,274—58 and 42 per cent respectively (DES, 1968, p. 17). In the sample the distributions were 57 and 43 per cent respectively (127 women, 95 men). The national distribution of women teachers in the particular year was 114,041 in primary schools, 58,835 in secondary schools—38 and 20 per cent respectively of all teachers (DES, 1968, p. 17). The sample contained 88 women in primary schools, 39 in secondary: 40 and 17 per cent respectively of the whole sample.

In discussing the final section Choppin accuses me of making statements about a population and not only the sample concerned. Throughout the paper I used cautionary phrases, for example 'the primary teacher here', 'in this sample', 'these teachers', 'peculiar to the samples used here', so that no one would be likely to mistake the shadow for the substance.

Though Choppin is entitled to his view on this, erroneous as I think it is, he really must not say that 'the sample was drawn entirely from the student body at one college' (p. 220) for it is untrue. He claims to have read the original (Cortis, 1968) and if he refers to it again he will find that the data collection took place in three colleges, not one, and that two were LEA controlled and one a voluntary college to allow for variations between the two types. Certainly the sample were younger teachers but on what possible grounds can Choppin say that they had 'little or no practical experience' and were '*prospective*' teachers (his italics)? *They were actually teaching and had been for two years and Choppin notes this in his preliminary remarks* (p. 219). It appears that in his intoxicating desire to polemicise he has read his own words but has not perceived their sense. It is as breathtaking in its audacity as in its carelessness.

Finally Choppin deplores my quoting Wall (1968) because he suggests I have not practised what he (Wall) preached in my 'lack of interest in the composition of this sample' (p. 220). The sample was close to national proportions of teachers in sex and employment distributions but obviously not in age. Wall's comments on 'the avoidance of error at the expense of reality' (p. 10) illustrate my position well. To sample teachers' behaviour over every age range (say up to 60 years) in terms of the extensive yet delicate level of instrumentation I used would be quite unreal, but it might (if it were possible to do, which it isn't) help to avoid error. I hoped I had made clear the limitations of my sample as I explained above. Very little work has been done in following-up the performance of teachers at this depth in Great Britain but the American literature is full of reports on samples far smaller and less representative than this. I should certainly have made more public the details of the sampling procedures, not least in my own interests, but this omission in no way justifies the mix of inaccurate facts and heady supposition that constitute Choppin's critique.

Nothing is in some ways more satisfying or more dangerous than to attempt to demolish others' work whether on sampling grounds or methodological grounds or whatever. The art of criticism has its own particular enticements. What the critic originally conceives as a daunting duty can so easily become a positive pleasure under the spell of seductive semantics. But with journal articles on empirical work it is particularly risky because they are an invariably condensed report of what actually took place and the form itself, by its very nature, tends to impose reticence on the writer (criticism of a full length book is an obviously different matter). I would have thought one's own research record needs to be very extensive indeed and oneself very eminent to judge a journal article in the manner here. In fact the eminent don't do it—at least not as the *raison d'etre* of a paper—presumably because *public* in-depth evaluation of any individual's report implies an omniscience few would care to lay claim to. Is this type of criticism then the 'new look' in educational research? I very much hope not.

References

Choppin, B. (1974). 'Interpreting research: sampling', *Educ. Res.*, **16**, 3, 218-21.

Cortis, G. A. (1973). 'An analysis of some differences between primary and secondary teachers', *Educ. Res.*, **15**, 2, 109-14.

Cortis, G. A. (1968). 'Predicting student performance in colleges of education', *Brit. J. Educ. Psychol.*, **38**, 115-22.

Department of Education and Science (1968). *Statistics of Education 1966, Vol. 4: Teachers*. London: HMSO.

Douglas, J. F. (1973). 'A study of streaming at a grammar school', *Educ. Res.*, **15**, 2, 140-3.

Quigley, H. (1973). 'The pre-reading vocabulary of children leaving nursery school', *Educ. Res.*, **16**, 1, 28-33.

Wall, W. D. (1968). 'The future of educational research' in Wall, W. D. and Husen, T., *Educational Research and Policy-Making*. Slough: NFER.

Bruce Choppin Replies:

Dr Cortis has found an error in my paper. My *one college* was an inaccurate contraction of *three constituent colleges of education of the University of Manchester School of Education* (Cortis, 1968). I apologize to all concerned for this unfortunate lapse.

The subject of my paper, which has drawn such bitter complaint, was sampling. Dr Cortis' sample was drawn from the student body of the three colleges concerned. While still students, these individuals were subjected to 'extensive tests of intelligence, attitudes and personality' (Cortis, 1973). It is surely not unreasonable to refer to this sample as one of 'prospective teachers . . . quite atypical of the teaching profession as a whole'.

The major question, however, is whether the Cortis (1973) article should be subject to critical review by other researchers—whether eminent or not. I wrote that Dr Cortis' article contained insufficient information about the sample to permit the thoughtful reader to properly evaluate the conclusions he seems to reach. I am still of that opinion.

The use of inferential statistics throughout, and the phrasing of the final section, lead the reader to believe that generalizations are intended. Dr. Cortis now tells us that he was careful to ensure 'that no one would mistake the shadow for the substance' but it would seem that I was not the only one to be dazzled by his mirage. Press reports quoted freely from his final section:

'Primary teachers are inclined to be rigid and insensitive about their work and rather shy in the classroom, a research programme has found. They are not satisfied with their work as secondary teachers and tend not to hold progressive attitudes.' (*The Teacher*, 2.3.73).

'Primary school teachers are more rigid and more conventional than those in secondary schools, according to a report in the latest edition of the National Foundation for Educational Research Journal.
The report says that secondary teachers are more sensitive, more satisfied with their work, and have more progressive educational attitudes than their primary counterparts.' (*Times Educational Supplement*, 2.3.73).

'Shy teachers face big problems in the classroom—particularly if they are trying to make their mark on very young children. Even children as young as five will notice and tend to misbehave.' (*Daily Mirror*, 27.2.73).

'Conventional images of primary and secondary teachers have been overturned by research carried out by Dr G. A. Cortis of Birmingham University School of Education.' (*Guardian*, 27.2.73).

It is hard to find evidence here that the results are not supposed to be extended to teachers outside the original sample. There are no clues for the reader that 'rigidity', 'sensitivity' and 'shyness' were derived from personality measures adminstered to college students and not to practising teachers.

The net result can hardly have strengthened the reputation of educational research. 'USELESS AND INSULTING' exclaimed *The Teacher's* headline; '. . . sensitive, intelligent teachers will be more annoyed than anyone that some researcher should have the gall to rush into the public prints with such a simplistic generalization'.

The reporting of one's findings is an important part of the whole research process. I am not alone in my concern about the quality of some published work. Individual educational researchers, have a responsibility to maintain professional standards, and criticism of published work must play a part in this.

MATHEMATICAL TALENT

Discovery, Description and Development

Edited by Julian C. Stanley, Daniel P. Keating, and Lynn H. Fox

More ought to be done, the editors assert, to help intellectually gifted young people move ahead educationally. They here report findings from their ground breaking Study of Mathematically and Scientifically Precocious Youth. The Study identified exceptionally talented seventh and eighth graders, studied their backgrounds and interests, and offered the most gifted of them ways to accelerate in mathematics either through special high school level courses or by having them take college courses for credit.

Julian Stanley gives historical background on gifted children and tells how the study originated. Daniel Keating describes data collected during the first year, Lynn Fox evaluates the methods used to facilitate the students' advancement, and Helen Astin discusses sex differences in the mathematics competition. The second part of the book includes an analysis of how well the group learned high school level mathematics, an evaluation of their social and emotional development, and a look at their adjustment to a college classroom situation. Each section of the book concludes with an overview and critique.

215 pages, Cloth £5.00 paper £1.50

Johns Hopkins University Press
2-4 Brook Street, London W1

Methuen announce

The British Journal of Teacher Education

EDITOR:
PROFESSOR EDGAR STONES,
University of Liverpool

This new journal, designed to keep all concerned with teacher education abreast of world-wide developments in a rapidly changing field, will examine the whole process of teacher education, the policies relating to its development, its institutional patterns, the changing content of initial preparation, in-service education and staff development. Two central objectives of the editorial policy of the journal are that it should be actively interested in current issues, and that, with the help of a panel of overseas correspondents, it should reflect developments in teacher education on an international scale.

The journal will appear three times a year, in January, April and October. First issue: January 1975.

Subscription rates: Individuals £5.50; Institutions £6.75; North America £6.75 (U.S. $18.00); airmail to all parts of the world £10.75 (U.S. $29.00); Single copies £2.25 (U.S. $6.00). All rates inclusive of postage.

Subscriptions to: Methuen & Co., Subscriptions Department, North Way, Andover, Hampshire SP10 5BE, England.

NFER

COMPUTER SERVICE

The National Foundation for Educational Research has now installed its own computer—an IBM 1130.

We anticipate a certain amount of spare time in the next year or so, and are prepared to make the computer available to Educational Research Workers at favourable rates.

Card-punching will also be undertaken, at competitive rates.

Details

A comprehensive range of programs is available. These include regression analysis, factor analysis (up to 100 variables), a general tabulation program, χ^2 analysis, and analysis of variance. In addition, there is a wide range of data processing programs for those researchers who need to handle a large amount of data.

Rates

Computer: £21 per hour.

Data Processing: £27 per 1,000 cards. Punched and Verified.

Specimen Costs:
Regression analysis of 8 variables on 40 subjects—£5: 15 tables from 2,600 subjects—£10.

All charges subject to 10% VAT.

Telephone Slough 28161, or write to

A. N. James,
NFER, The Mere, Upton Park, Slough, Bucks.

chambers

●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●●

Young Set Dictionaries

Amy L. Brown, John Downing and John Sceats

for each primary school child – the right dictionary

A unique set of graded dictionaries. Illustrations, which constitute *Dictionary One*, are gradually phased out in *Two, Three* and *Four*. In these, however, the definitions are expressed in the language of children—simple, colloquial and to the point.

'conceivably the most useful and usable children's dictionaries ever produced.' ***The Times Educational Supplement***

Dictionary One	45p	Dictionary Three	50p
Dictionary Two	45p	Dictionary Four	60p

Inspection copies from
Chambers 11 Thistle Street Edinburgh EH2 1DG

Book Reviews

Education for Liberation

by Adam Curle

Tavistock Publications, 1973, 144 pp., Hardback £2.50

Adam Curle's delightful book is a revolutionary tract containing practical guidance in the use of non-violent techniques for subverting society. The school curriculum of the future will rest heavily on case-studies of successful non-violent revolutions. But effective subversion is not simply a technical and organizational affair: it is the outcome of greatly extended 'awareness-identity', and the downfall of 'belonging-identity'. Adam Curle has examined these concepts in greater depth in an earlier book, *Mystics and Militants* (1972). In this new book he deals more directly with the part that education must play in the revolutionary process. He is not a de-schooler. The school is potentially a powerful revolutionary device. Schools and their teachers must, of course, be transformed; but Winchester and Eton—collapsed into 'Winton'—might remain, one feels, almost precisely as they are.

It is true that even Winton serves the system of competitive, exploitative materialism, although in the quality of its teachers and the many-sidedness of its pupils it seems already qualified for a place in the counter-system. Aside from Winton, schools deal only in death; and Adam Curle examines briefly some of the varieties of scholastic enslavement in both the developed and developing worlds. He sees the quality of teachers as crucial to the new reformation, and he would disestablish the teaching profession (and, indeed, following Illich, all other professions). The teacher will change his relationship to 'his' knowledge, but there is no chance that he will do so if he is trained in a University Department of Education. 'Of all types of education the training of teachers is perhaps the worst.' In his period of training the student teacher 'merely learns how to control spontaneity, crush initiative, instil fear or contempt, enforce passivity, impose meaningless regulations and senseless standards, and to reduce to mindlessness the whole process of education' (p. 75). Teachers will be chosen in mature years by a commission skilled in identifying wisdom, patience, sincerity, warmth, and inner coherence. Selection for teaching will be one of the highest honours that society can confer.

This is a book of recantation. The export of our educational system and values to developing countries promotes strong belonging-identities where formerly they were often commendably weak; and it has a dubious relationship to economic growth.

['Having written three books and countless shorter pieces on this topic, I have now reached the point where I believe that most of what I wrote was wrong' (p. 114).] But growth in any case is a false god and its pursuit a sign of lack of awareness. Growth in advanced countries is simply excess, and these excesses are not only unnecessary but anti-developmental. In the future men and women will learn to need less. It is then that the counter-system will prevail.

This book is a counter-cultural manifesto. It is written with transparent sincerity, out of deep learning and a wide experience of affairs. It presents no new facts or theories. It is a plea from a deeply humane man that humanity should at last be human. The deepest irony of the book is that it was written deep in the woods of Maine and its author is now a senior member of one of our technological universities. This will be a sterner test of humane values than Ghana or Bangladesh. One can only hope that the author perfected his tactics of subversion before he accepted the job.

FRANK MUSGROVE

The World Year Book of Education, 1974, Education and Rural Development

edited by James R. Sheffield and Philip Foster

Evans Brothers, 1973, 417 pp., Hardback £6·00

In these times of crisis in the industrialized world it is salutary for all of us to have kept before our eyes the fact that two-thirds of the world's people still live in countries where the average per capita income is under $250 per year, where the gap between rural and urban living standards appears to be widening rather than closing, and where the absolute number of unschooled and illiterate adults is increasing rather than diminishing. For all of us concerned with the future of mankind, an evaluation of the influence of education and schooling upon social and economic life is of great importance, and for those educators in particular who are working in developing countries or are involved in the preparation of teachers and administrators for those countries there are some crucial questions about the value of certain approaches to education upon the answers to which depends the possibility of success for their efforts.

This volume of the *WorldYear Book of Education* will be invaluable to the educator concerned with the developing world. Both from the theoretical and from the practical point of view, the Year Book is a rich source of information and ideas about the main problems of education for rural development. The twenty major papers presented in it are grouped into two sections. In Part I the contributors deal with critical problems such as the varying perceptions of the aims and functions of education in the developing countries, the social, economic and cultural perspectives of rural transformation, the relationship between formal schooling and non-formal agencies of education, the use of the mass media and of alternative methods of education, and so on. Part II consists of case studies illustrative of various approaches to education for rural development in many parts of the world. Radical models of rural transformation, such as those of China, Cuba, and Tanzania, based on explicit ideological premises are described, but, owing to a lack of directly relevant factual information, little in the way of a precise evaluation of the results has so far been able to be made. On the other hand, from several countries where there has been a less explicitly radical orientation, countries such as Burma, India and Malaysia, we have some detailed studies that carefully tackle some very specific problems and provide hard empirical data that really do help to clarify some development questions of great importance. An example of this kind of study is Bruce R. Harker's paper, which, using the technique of path analysis, examines the relationship of education to innovatory behaviour among Japanese farmers.

The editors of the Year Book state that, realizing as they did the vastness of the theme, they made no attempt to deal with all the main problems or to cover geographically all the areas in the world where rural education was important. Instead, they sought papers from authors who could make significant contributions either to the theoretical analysis of some of the outstanding issues in development education or to the presentation of empirical data relevant to such issues. An interesting example of the former kind of study is C. Arnold Anderson's typology of curriculum content and his analysis of the relationship of education to the dynamics of agricultural progress. D. P. Chaudhri's succinct demonstration of the relationship between the level of education and the level of agricultural productivity in 256 districts of India is a good example of the latter.

It is hardly to be expected that in a volume of this kind, presenting work by a wide range of contributors, major generalizations would emerge that would command general agreement. Nevertheless the editors do find some that are of considerable interest. One of these may be mentioned here because of its almost universal topicality at the present time, and that is the conclusion that non-formal agencies of education can be regarded as a substitute for formal schooling only in certain circumstances; they are best viewed as being complementary to formal schooling. We are given the salutary warning not to let non-formal (or informal) education be thought of as a panacea for all the educational ills of developing countries nor to let it become the fad of the 1970s. The real task of research workers into this aspect of development education is to assess the respective advantages and limitations of various types of educational programmes in the schools and outside of them. This task, say the editors, will involve 'careful, comparative research—not the proclamation of new orthodoxies'. Like many of its predecessors, this Year Book itself should do much to encourage needed research, by focusing attention upon crucial issues, by demonstrating the main gaps in our knowledge in important areas of education, and by indicating some of the more promising fields of enquiry and methods of research.

It should be noted here that this is the final edition of the World Year Book of Education, the editorial board having regretfully decided that, with the proliferation of educational research journals of a more specialized nature and indeed with the increasing volume and specialization of research itself, it had become extremely difficult to continue to publish the kind of comprehensive survey that had been pioneered by the Year Book. The contribution made by the Year Book to educational studies since it was founded in 1932 has been carefully analysed by Brian Holmes in a postscript to this volume.

G. W. PARKYN

Tinker, Tailor . . . The Myth of Cultural Deprivation

Edited by Nell Keddie

Penguin Education, 1973, 150 pp., Paperback 45p.

Theories of cultural deprivation as explanations of scholastic failure are dangerous because they entail a tacit acceptance of the mental deficiencies of lower-class children whilst diverting attention from the cultural life of the social situations to which they belong. Furthermore, as Labov points out, such theories are likely to lead inexorably to the hypotheses of genetic inferiority that they were originally intended to preclude. This book resists the whole intellectual framework of deprivation theories and will be welcomed with a sense of relief by those attempting to reach beyond this crude positivist approach in social sciences. It consists of seven previously published papers which offer a coherent analytical framework for the explanation of the poor educational achievements of lower-class groups; a framework, furthermore, that demands neither genetic hypotheses nor a resort to deprivation-induced deficiencies.

A book of this size, in spite of its scope, is inevitably highly selective in content. On the other hand, as Nell Keddie complains, there is a very real paucity of research, especially in Britain, that has conceptualized the problems realistically. Three of the chapters, then, relate to black ghettos and slum-schools in the USA. Labov, in a now classic thesis (The Logic of Non-Standard English), challenges the illusion of verbal deprivation, claiming that the traditional research approach, hindered by preconceptions of verbal ability, has been unable to tell us anything about the language of the groups studied. Hence the conclusion has been that their language is restricted, crude and non-logical, with little capacity for creative and abstract modes of expression. Labov, in his research with Harlem black children, shows how the prior acceptance of negro language as an adequate, coherent and functional system, guided a research strategy that yielded rich fruits in terms of linguistic data. Penetrating traditional barriers reveals that negro children, labelled as linguistically deprived, in fact participate fully in a highly verbal culture, and hear more well-formed sentences than middle-class children whilst having the same vocabulary, the same capacity for conceptual learning and the same logic as any other user of English. Labov shows that negro language is a separate system, not a deficient one; it differs in systematic, rule-governed ways from standard English whilst expressing equivalent logical content.

Jane Torrey is concerned with literacy programmes and argues that the main impact of negro dialect on education has not arisen from any structural difference but rather from its function as a low-status stigma. It becomes branded as wrong and something to be stamped out. She appeals for the academic world to broaden its cultural base beyond the traditional subjects, methods and media, and, in relation to minority cultures, for the relegation of standard English as a second dialect for special purposes. These questions are placed in the context of school organization by Estelle Fuchs. She illustrates, from recorded accounts of experience, how newly recruited teachers quickly become habituated to the institutional imperatives and accept the prevailing rationale about student failure.

Literacy is also the concern of Neil Postman who challenges the belief that literacy is the richest source of aesthetic experience. The teaching of reading he sees as a sinister political activity; its major consequence is the student's accessibility to political and historical myth; less than one per cent of them will actually use it in pursuit of the lofty goals always projected. Postman also embraces the McLuhanite prediction that the new electronic media will disrupt social and political ideas, awaken new forms of sensibility and make reading a skill of diminishing importance. This highly provocative paper, published three years ago, has not yet evoked the wide critical reaction one might have expected. One obvious criticism made by Nell Keddie is that the new media are just as amenable to the preservation of myth if the power relationships across them remain unchanged.

The remaining three chapters are anthropological. Margaret Mead's paper, first published thirty years ago, has acute relevance to contemporary issues. She points out that education is a cultural process, and then examines the process in a primitive perspective; 'We have no way of knowing how often in the course of human history the idea of truth as a revelation to or possession of some one group (which thereby gained the right to consider itself superior to all who lacked this revelation) may have appeared.' This 'superior' group however, then engages in active proselytizing. Education becomes an adjunct to those in

Education Books from Methuen

What School Is For

GABRIEL CHANAN and
LINDA GILCHRIST

This remarkable and radical book cuts through the rhetoric and confusion in educational theory to confront the basic dilemmas with which every teacher is faced in creating a productive relationship with his pupils. The authors, both experienced teachers, show that the culture of the pupils and the culture of the school do not have to be incompatible, but that their synthesis demands nothing less than the reconstitution of school as a workshop for the creation of a new culture. £1.30 paperback 65p

Contemporary Research in the Sociology of Education

Edited by JOHN EGGLESTON

A representative selection of the most important recent research in the sociology of education, ranging from 'mainstream' studies of educational opportunity to newly developed interactionist and phenomenological approaches to the school, classroom and curriculum. £4.80 University Paperback £1.95

power and the business of those who teach rather than those who learn. This entails a deliberate violence to the developing personality of other people and the emphasis is shifted from one of growth and learning to one of dictation and forced acceptance of clichés and points of view.

In answering the problems of defining 'mentally normal', Thomas Gladwin illustrates the contrasting style of intellect that can arise in different cultures. He compares the ocean navigation of Trukese Islanders with that of Western sailors. The cognitive strategy of the latter is to construct an overall plan that is implemented piecemeal. The Trukese, however, implement a continuous series of improvisations on an *ad hoc* basis moving dialectically towards the end-point (a tiny island some hundred miles away). In a more systematic investigation of disease terminology among the Subanun, Charles Frake provides a significant demonstration of a long standing postulate, that the meaning of a linguistic form is part of the total context in which the form is used. This has been illustrated in other ways, for example by Brown and Beluggi in the acquisition of syntax. What is important in Frake's study is the basis of cognitive differentiation between groups. Thus the number of levels of 'contrast' in a particular category (in this case disease) corresponds with the predominance of that category in a wide variety of socio-cultural settings. In a neighbouring culture an alternative category (say botanical taxonomy) may be predominant; but any comparisons from the point of view of the Subanun would always lead to the conclusion that this other culture is deficient in powers of abstract thinking. However, as Frake explains, the disparity must be explained by the cultural uses of respective systems 'and not by appeal to differences in the cognitive powers of individual minds'.

All these contributions, then, offer alternative ways of conceptualizing the problems that cultural deprivation is supposed to explain. But as Nell Keddie explains in her critical introduction, they also pose new questions demanding new research orientations. One of these must be the relating of these cultural-anthropological concepts to our own social structure. The implications of a hierarchical division of labour and attendant subculturalization are not often grasped in egalitarian educational writings.

K. RICHARDSON AND M. CALNAN

Second Handbook of Research on Teaching
edited by R. M. W. Travers
Rand McNally, Chicago, 1973, 1,400 pp.,
Hardback £12·50

Gage's *Handbook of Research on Teaching* must be one of the most frequently consulted books on the shelf of any educational researcher. Its advocacy of particular observational and experimental approaches has exerted a considerable influence on the methodology of research during the last decade, and many of the review chapters are exemplars of their kind. The success and importance of Gage's *Handbook*, and the tremendous activity of recent years, has prompted the American Educational Research Association to produce a *Second Handbook*. This is, however, no mere update of the first, but a completely new and complementary collection of chapters prepared by some of the most eminent experts in their fields.

Gage's *Handbook* was put together round a tidy framework of central variables (behaviour and characteristics of teachers), relevant variables (social interaction in the classroom and the social background of teaching) and site variables (grade level and subject matter), but research on teaching is not now easily contained within this neat scheme, and in the *Second Handbook* we get an altogether freer organization of information and ideas. In all there are forty-two chapters grouped into four sections: introduction, methods and techniques of research and development, research on special problems of teaching, and research on teaching of school subjects. The introductory section contains chapters on contemporary models of teaching, theory construction, the impact of research on teaching, and social and political influences on educational research; the methods section is mainly concerned with observational methods and the analysis of qualitative data; the special problems treated include the teaching of special groups (the mentally retarded, the emotionally disturbed and the gifted), the teaching of particular skills, and teaching at different levels; and, in the last section, reviews of research on particular subjects covered in the first handbook are brought up-to-date and other areas like physical education and business subjects are added. The quality of the chapters is generally as high as the quality of the material would permit.

The contributors evidently did not find their tasks easy. In the preface, Travers, the editor, writes: 'a majority of the authors, after they had completed a review of the relevant literature, expressed disappointment about the lack of substantive research in the area in which they had agreed to write a chapter'. A common response was 'to begin to consider why the research had not been productive and how a productive program of research should be planned'. Here, I think, we have the unifying feature of an otherwise diverse and disparate collection of writings: most of the contributors seem to be asking, either implicitly or explicitly, what has been wrong with research on teaching and how can it be improved?

Among the problems mentioned are: lack of an adequate theory (Getzels and Dillon; Dessart and Frandsen; Nixon and Locke), insufficient regard for the specificity of teaching (Trent and Cohen; Beller), too much emphasis on outcomes and not enough on teaching activities, teaching situations and the learner (Bidwell; Birkmaier), not enough attention given to the environment (Green, Bakan, McMillan and Lezotte) and that research is ofte short-term and carried out mainly for higher degrees (Shulman and Tamir; West).

Some feel that more careful attention to existing procedures might suffice to bring about the necessary improvement. Thus, Trent and Cohen, taking up where Wallen and Travers left off in 1963, conclude that studies of teaching methods have 'yielded no clear-cut evidence of the superiority of any one approach' and suggest that this may partly reflect problems in definition, design and measurement. Shulman and Tamir, reviewing research on teaching natural sciences, criticize recent experimental studies as being 'short-term, poorly analysed statistically and ambiguously reported'. Barber is able to identify at least nine 'pitfalls' in conventional research (for example, 'the investigator loose protocol effect') and concludes 'there are so many pitfalls in any one experimental study that we should not take any one study too seriously'. His solution is for extensive replication. But, of course, it is the problem of replication that makes it so difficult to study education scientifically.

But others, like Eisner who advocates a 'critical clinical' approach and Shaver and Larkins who recommend the development of a 'classroom ethnography', suggest that a radical change in perspective is required to unravel the subtleties of teaching. How such a systematic study of teaching in its natural setting might be attempted is suggested by Rosenshine and Furst in a notably good chapter on the use of direct observation to study teaching.

Richard Light in his perceptive chapter on the treatment of qualitative data argues that the basic problem in the study of teaching is one of measurement; by forcing human behaviour to fit common scaling techniques when we have no reason to believe that they are appropriate we are committing 'a double non science': 'the model chosen in the first place may be arbitrary, and even after the model has been chosen the variables are scaled or measured in an arbitrary manner'. He goes on to show that not only are the techniques available to analyse qualitative data, but also that they are relatively simple to apply. He also offers a tentative approach to the much more dificult problem of deciding what kinds of model could have generated the observed data. But this is really the province of Snow who offers an interesting and comprehensive chapter on theory construction for research on teaching. Together these two chapters could exert a considerable influence on research methodology during the coming years.

The 1,400 pages of the *Second Handbook of Research on Teaching* are a massive consolidation of the present state of American knowledge on research on teaching. Like its predecessor, it will probably become an indispensable work of reference for all who are interested in the field. But the essence of effective teaching still seems as elusive as ever.

ALAN SMITHERS

The Education of Teachers in Britain

edited by Donald E. Lomax
John Wiley and Sons, 1973, 467 pp.,
Hardback £5·50

This book was published early in 1973; most of the writing was carried out before the end of 1971. The James Report was published early in 1972 and the White Paper, 'Education: a framework for expansion', was published in December 1972. Some of the authors introduced modifications to their chapters after one or both of these documents appeared. But in effect the book represents a stocktaking at a turning-point in the history of teacher-education in this country: it provides a series of critical reviews of current practice and from these a number of far-reaching and imaginative pointers to future developments. Much of it will no doubt in another decade be respected as a monument to the achievements and the frustrations of the three decades following the McNair Report; but much of it, as a result of Dr. Lomax's recruitment and briefing of his team, will be recognized as perceptive, prescient and practical.

It is impossible in the space available to review the whole of a volume of this nature. Readers of this Journal are likely to be particularly interested in the research implications of the various contributions, but even here it would take a very substantial paper to trace the full implications for research of the whole volume. The selection of material for comment is therefore bound to reflect the preoccupations and prejudices of the reviewer.

It is natural to turn first to Part 7, on 'Research in teacher-education'. Dr. Louis Cohen surveys the material on the social and academic characteristics of student teachers and on their personality and professional attitudes. He and Professor McLeish then devote two chapters to the highly significant question of the impact of college communities on the students who pass through them. They refer to the evidence on changes towards more 'radical' and more 'progressive' professional views and on the influence of close staff/student contact and of subject-group and peer-group sub-cultures. They disentangle to some extent the very varied influences operating even in a smallish college community and point to the need for longitudinal studies which will trace the changes of attitude and outlook that take place in the early years of a teacher's career.

Dr. John Garner of Armidale overlaps these chapters in the final contribution and extends their thinking in two directions. Firstly he commends the increased attention now being given to the influence of specific types of teaching behaviour on specific aspects of classroom learning, and carried out by specific types of children, referring particularly to Soar's findings that (among other things) children's creativity and reading skill are differentially influenced by a preponderance of non-directive behaviour by the teachers in charge of them. He does not, however, draw out the implications of this evidence, for the techniques by which *teacher-educators* may seek to develop particular traits in particular types of *student-teachers*. Some of the work of Professor A. Prost of Orleans is relevant here. Second, Dr. Garner underlines the point that every choice of aims and objectives in teaching is based on—or implies—a value-decision.

It is a merit of this book that the Editor has found space for wider implications of this principle to be opened up in Part 4 of the book, by Dr. David Aspin. This author makes, in the course of his argument, some perceptive criticisms of the thinking of the James Committee on 'breadth' of education. In view, however, of the trend of argument in Part 7 already noted, the most significant comment of Dr. Aspin's is probably that 'the initial questions raised in any discussion concerned with the place of values in teacher-education (will concern) the nature of value, the complex inter-relationships of values, value sets and value systems, *and—perhaps the crucial issues—the justification of our value judgments*.' (Italics not in original.) There is evidently some incipient convergence of views on the central importance, in teacher-education, of the exploration not merely of concrete and over-simplified behavioural objectives, nor only of higher-order aims involving complex psychological/ethical concepts, but also of the deeper value assumptions and commitments, whether professional or personal, which young adults may discover and articulate in their own minds; and in addition of longitudinal studies: which could among other things attempt, however rashly, to trace the sources of (say) 40-year-old teachers of distinguished reputation, of their specific and characteristic qualities, whether conservative or radical, dedicated or detached, authoritarian or liberal, or whatever.

In terms of research on teacher-education, then, the book is both judicious and suggestive. It is to be earnestly hoped that the Social Science Research Council and the DES will, despite the present financial stringency, find the money to sponsor research on an appropriate scale. The Committee for Research into Teacher Education (CRITE), a joint body composed of representatives of the Association of Teachers in Colleges and Departments of Education (ATCDE), the Universities Council on the Education of Teachers (UCET) and the National Foundation for Educational Research (NFER), together with observers from the DES, the Schools Council and SSRC, has embarked on the task of promoting a climate of opinion favourable to such investigation and planning a strategy of study. What is needed now is a programme of investigation as wide-ranging as that of the Schools Council for the schools and as adequately supported, not only by the DES but by SSRC. Momentous changes are now being planned for teacher-education over the next decade: in no sphere of education has it become more important for an adequate machinery to be established for monitoring the fresh developments—which cannot be satisfactorily carried out without systematic studies of present practice.

The book itself is excellently produced. There are one or two odd failures, such as the omission of the references given in Chapter 11 from the bibliography provided at the end of the book. Nevertheless a well-planned exercise has been executed at a well-chosen moment.

K. G. COLLIER

Physical Education and Sport: Sociological and Cultural Perspectives

edited by J. A. Mangan

Basil Blackwell, 1973, 238 pp., Hardback £3·00

Instead of writing this review, I could be watching on television, either test cricket, tennis from Wimbledon or World Cup football. While millions of people throughout the world will be watching one of these sports, it is unlikely that any sociologists will consider either the playing or the watching to be worthy of serious sociological investigation. It is the lack of interest by researchers into the sociological analysis of sport and physical education that has, in part, prompted Mangan to edit a set of 20 introductory readings into 'the sociological and cultural perspectives of physical education and sport'.

The text is divided into three quite independent sections. In general, the readings in the first section examine relationships between student characteristics and performance in physical education and sports. They indicate our limited understanding of how participation in team games and other physical activities influences student characteristics such as self-concept, attitudes to school, peer group relations, and academic performance. The readings also suggest that there are likely to be interesting relations between the organizational structure of schools, student involvement in sports and student involvement inside the classroom. For example, if a school changes from placing stress on team sports in which only a sporting élite is involved to a school that emphasises physical pursuits requiring involvement at an individual level, then quite different student attitudes to school might result. However, before meaningful research can be undertaken into the sociology of physical education and sport in schools, we need to develop theoretical frameworks to guide the studies. The first section of the text is a move in the direction of suggesting some possible frameworks.

The second section of readings investigates the status of sport in society and analyses some of the relations between the concepts of leisure, play, sport and society. Although society's involvement in physical education and sports has continued to increase, there has been little real attempt by sociologists to include an analysis of sports in their construction of conceptual frameworks of society. As a number of readings suggest, the academic status of sports sociology is not particularly high. In part, the low status is related to the oversimplified nature of the conceptual frameworks that are used in investigations and the lack of sophistication in the methodological and statistical approaches to problems. The same criticisms that may be attached to sports sociology in general are reflected in the present set of readings as they do not attempt to build up theoretical frameworks from which meaningful research may be generated.

The relation between sport and physical education and personal well being is examined in the final set of readings. A number of biosocial articles analyse the link between performance in physical education and the health of individuals. There are suggestions that for health reasons there is perhaps a need to introduce school students to those physical activities that may be pursued for many years after leaving school rather than concentrating on team games which tend to have minimal participation transfer after leaving school. The health reasons that are provided in support of multiple choice programme of physical education links with the social reasons that are provided in the first section of the text.

Mangan is very much aware of the limitations of the research that has been undertaken so far in the sociology of sports and physical education, and indeed a number of the present readings reflect these limitations. But the text does provide an interesting contribution for those who research in the area of the sociology of schools. Its major contribution is the research questions that it generates and these questions should provide many students with starting points for future investigations.

KEVIN MARJORIBANKS

Education 3-13

A journal for Primary Education

In the October issue

Whose school?: 'heads.' win or lose
A school without a head
To hell with registers!
Open schooling: an introduction
Open space – open classroom?
Reading: six to nine

It may be ordered direct from the Publishers

In Britain

The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow G64 2PW

Annual subscription £1.50 including postage

In the U.S.A. and Canada

APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

The Study of Education and Art

edited by Dick Field and John Newick
Routledge and Kegan Paul, 1973, 244 pp.,
Hardback £3·80, Paperback £1·90

This work discusses, in short compass, a wide range of important themes in education and the arts in general, not just visual art. The authors are all closely associated with the University of London Institute of Education and one detects, as a consequence of prior mutual influences and disagreements, a public rehearsal of previously established positions; this gives the book a certain degree of coherence both in the employment of a consistent terminology and a reasonably uniform share-out of the individual themes discussed. Despite this, some glaring contrasts of outlook and approach emerge, not merely as matter of style, but in level and quality of argument.

Nevertheless, the work is well worth studying for the majority of its contributions, several of which, though necessarily brief, contrive excellent summaries of their subjects. One hopes that its steep price (which at least ensures a high standard of printing and binding) will not deter students of education, who should certainly read it. Researchers in this field will also find it of interest if only because it throws out many indications of themes worth pursuing, most notably an inconclusive discussion on the question 'Is is necessary to make art in order to teach art?', and the related question of what the arts are in education and how they are to be distinguished, e.g.: 'performing' versus 'making' (or interpretive versus creative).

It is necessary to mention each contributor singly, one cannot avoid partiality and prejudice, in view of the contrasts previously mentioned. Rosemary Gordon and Mel Marshak, discussing quite different topics, share a common Jungian starting point, much of their argument being anti-materialist and mystical, even bordering on the magical, and both share an assumption of the supreme importance of religion. Their styles, replete with portentous phrases and neologisms, are curiously similar, and difficult to follow in printed form; nevertheless, they pursue important issues, for instance, Marshak's view that the incorporation of psychological insights into educational philosophy is sinister because it has more efficiently manipulated individuals to be subordinate to the social order. Like much in her essay this view, proceeding from questionable assertions and reaching irrational conclusions, seems provocatively advanced and yet is also obscurantist and reactionary.

John Newick and Leslie R. Perry provide a refreshing contrast in two clear and well argued pieces. Newick attempts a dynamic view of the inter-relationship between environmental and personal and social factors, which unfortunately goes no further than unsubstantiated assertions; but he succinctly adumbrates the questions about his function which every good teacher should ask himself. He defines the teacher's role culturally, psychologically and morally, thus leading into a discussion of values, of the nature of the sharing of experience and the responsibilities of the teacher and the school. Perry demolishes what he calls 'the cult of leisure', and discusses with much common sense the nature and value of art teaching, especially in relation to art history, including a stimulating discussion of adventurousness compared to creativity. He also has sensible things to say on art as communication and the validity of aesthetic judgement.

Dick Field clears up some confusion in his analysis of 'child art', and thus is mostly concerned with pictorial art (design is mentioned but the subject is not pursued and the argument is left unresolved). He scores points off Herbert Read (an easy target) but overlooks his real significance—the social and epistemological implications of his (admittedly vaguely defined) ideas. Field's articles poses innumerable pertinent questions but scarcely attempts to answer them, except by non-sequiturs.

The clarity of Louis Arnaud Reid's contribution makes it a pleasure to read, except for a few terse passages on problems of aesthetics such as content and form, where the space he allows himself is inadequate to carry any sustained argument. In the main it is an extremely sensible account of the necessity and the means for aesthetic education and its value. He fruitfully uses the concept of 'tacit knowledge' to unify his ideas, which are of highly practical value to the teacher willing to utilize them. This chapter, and Sonia Rouve's on teaching art history, are for me the highlights of the book. In the process of exploding the 'cult of objectivity' in art history, she stresses the importance of having a conscious philosophical standpoint in teaching art history, and advocates a pupil-centred, structured approach. She has wise words and good practical advice to offer on a methodology for working in depth in art history, for instance disregarding conventional causative and chronological systems, all rooted in a strong philosophical base.

In its comparatively short length this book covers a great deal of ground, some of it treated summarily, but providing a fertile bed for further cultivation; and despite many points of conflict, for instance on the uniqueness of art experiences compared to those encountered in, say, science, there are several general conclusions on which there is near unanimity. The most important, perhaps, is that of the vital necessity in education for both attempting to make, and experiencing, works of art. (A final personal caveat: in this connection I was sorry not to see the role of museums further explored.)

LIONEL A. BURMAN

Teaching Reading Through Motor Learning

by Dorothy D. Sullivan and James H. Humphrey

Charles C. Thomas Publishers, 1973, 149 pp.,
Hardback $7·95

The title is somewhat misleading. Nowhere is it suggested in the book itself that reading can be directly or exclusively taught through 'motor learning'. The book represents a physical educationist's perspective upon the teaching of reading. The activities—mostly games—are intended mainly as ancillary or enrichment components of reading teaching programmes.

Two distinctly different ways in which motor learning impinge upon the reading process are described. This perhaps reflects the different specialist interests of the two authors.

On the one hand there is the development of fine psycho-motor functioning, as an intrinsic part of the reading process, and the more general aspects of motor and physical maturation which have psychological implications for general educational progress. This is discussed in a sound manner and some current US perceptual-motor therapy schemes are described. Some of the murkier controversies, such as that concerning laterality, are summarized. In this respect the book provides a useful introductory text which nevertheless takes a specialized approach. In a more general sense it serves to re-assure those who might be mystified or alarmed by physical activities and movement sessions in Infant and Junior schools that some eventual 'academic' benefit and purpose is served thereby. It may even go so far as to guide teachers with a remedial interest in devising physical activities which will serve teaching interests in a constructive, if indirect, manner.

The same goes for the second concern, which is with the use of games and action-based teaching materials as a motivating context for learning skills and drills. The skills so learned include the development of comprehension skills on listening —acting—reading of stories, particularly narratives that involve games for pupils to play such as those found in the *Read and Play Series* (published by Muller, 1965). More simple drills such as sound-discrimination and letter recognition based on semi-competitive team games are also described, and the final chapter suggests how these games can be used as a context for diagnostic observation.

Practising teachers may find ways of enriching their teaching in this book; many may feel that they are already applying a 'motor learning approach' but will perhaps be able to refresh their understanding of the theoretical purpose behind many games and physical activities. Schools following a fairly conventional graded-reader based approach to the initial teaching of reading in isolation might benefit by diversifying their reading curriculum to include a motor learning approach. *Teaching Reading Through Motor Learning* gives some positive suggestions as to how this could be done by introducing new activities or by finding new purposes for established ones.

DENIS VINCENT

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £2·75 post free. Overseas (excluding USA and Canada*): £3·20 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: 95p post free. Overseas*: £1·10 post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA and Canada: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Enquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Berks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the text, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or
2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© National Foundation for Educational Research in England and Wales, 1974

Designed by H. A. Senter

Printed by King, Thorne & Stace Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

25 SEP 1975

In this issue

Classifying educational objectives
Team teaching in FE
Interdisciplinary education

A/c
Docket & issue
to the librarian
24/9/75

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd.

EDUCATIONAL RESEARCH

Volume 17 Number 2
February 1975

Contents

Stimulus material in 'O' Level History
C. M. Ferguson and S. Garrett 83

Adding and averaging grades *J. Nisbet* 95

Attitude measurement: a critique of some recent research *P. L. Gardner* 101

Attitudes of Nigerian secondary school children to programmed instruction in Geography
P. O. Okunrotifa 110

The role of the physical education teacher *L. B. Hendry* 115

The impact of teacher life changes and performance on student dropouts *E. Carranza* 122

Adjustment to school *B. Thompson* 128

The prevalence of behaviour problems among children in care *B. A. Akhurst* 137

Junior school pupils' rejection of school library books *G. Fenwick* 143

Book Reviews 151

D 942(3)
L 28/7/75
E

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 17 Number 2
February 1975

Stimulus Material in 'O' Level History

Carolyn M. Ferguson and Sean Garrett, University of London School Examinations Department

Summary

This paper describes three investigations conducted by the University of London School Examinations Department into the possible use of stimulus questions, containing original historical material such as cartoons, maps, photographs and statistics, in the 'O' level history examination.

The results suggest that the new type of question works as well as or better than the conventional essay type of question. The various types of stimulus material seem to act differently. In general, candidates prefer and perform best in the verbal type of question. Different patterns of preference and performance emerge for boys and girls.

Against expectation, candidates were not afraid of the statistics question and both boys and girls scored equally highly on this. Girls preferred the cartoons and boys the map questions; photographs were not popular with either.

A comparison of the performance in the pretests with the GCE examinations shows that the stimulus questions with their short, structured sub-questions seem to favour the boys.

Aim of the research

It is generally agreed that effective examinations are related to the objectives for which the subject is taught. Modern history teaching attempts to make the student familiar with historical sources and materials. He must understand a given body of historical evidence and select and organize this information and be able to develop a historical argument in the light of the given facts. History examining must reflect these objectives and so provide questions that involve understanding, comparison, association and extrapolation rather than mere factual recall. With the use of stimulus questions, that is questions that contain original and other historical material such as maps, photographs, cartoons, tables and documentary extracts, we are able to provide an examination in which the skills of the historian can be employed with the raw materials of history.

The University of London School Examinations Department has recently decided to include questions based on stimulus material in the examination of 'O' level History Syllabus B. This is the most popular of the 'O' level history syllabuses and covers English and European History from 1760-1955.

Such a decision affects many thousands of students—in June 1972, 18,886 students in the United Kingdom took the examination for this syllabus, of whom 17,774 were from schools, the remainder being either from colleges of further education or entering as private candidates. Clearly a decision affecting so many could be made only after careful research.

Research was, in fact, carried out on three separate occasions, in 1970, 1971 and 1972. The size of the sample and the research methodology used varied in each case. There were, however, two basic aims common to all three projects:

(i) To test on GCE students a variety of stimulus-based questions differing in type and historical content in order to find out about the relative popularity of such questions and to compare students' performance with the essay-type questions of the existing examination.

(ii) To give experience to examiners in the construction and marking of this new type of question.

The first test was held in 1970 at a maintained boys' school in the Midlands. This was run as a feasibility exercise to find out if students would understand this type of examination question and would be able to answer it under examination conditions with a time limitation. The results of this small scale test showed quite convincingly that students found no difficulty in understanding stimulus material questions and could answer them under normal examination conditions.

In May 1971 and May 1972 further tests took place. These were on a larger scale. In 1971 seven schools participated with a total of 477 candidates; in 1972 four schools participated with 219 candidates.

This article deals with these latter two projects. The scope of each is summarized briefly and the results discussed. Because it threw up more interesting results, the 1972 project is discussed in greater detail than that of 1971.

The 1971 project

Two papers were drawn up, each comprising a section of either written or visual stimulus material questions (Section A) and a section of essay questions (Section B). In Paper I all the stimulus questions were based on documentary extracts; in Paper II all the stimulus questions were based on maps. In each case there were six stimulus questions, three based on English topics, three on European. The stimulus questions in the two papers did not cover the same topics, but in both cases the questions were chosen for their central importance in the syllabus so that performance could reflect the students' reaction to the particular question type rather than the historical content of the question. The second section in each paper comprised 30 essay questions. These also differed between the two papers. Students were asked to do three questions—one stimulus question, two essay questions. At each of the participating schools the two papers were handed out alternately to the pupils sitting in alphabetical order by name.

Table 1 gives a comparison of the candidates' performance in the pre-test with the GCE mark for four of the schools in the sample.

An analysis of variance on this data shows that there are significant differences in the performance for the two types of pre-test both within and between the four schools. This is not so for the candidates' performance on the GCE examinations; all the candidates may be considered to be of equivalent standard. Therefore it seems likely that differences in pre-test performance are due to the nature of the examination.

Clearly there is a relationship between candidates' performance on the pre-test and the GCE examination—the level of correlation may be as high as ·78. Interestingly, the levels of correlation appear very different for the only mixed school in the sample, School 2. This was also the school where the average marks of the two pre-tests differed by 21·6 per cent.

The results of this research were encouraging. They showed that the standard of student response differed for different types of stimulus material and it was thought necessary to investigate a wider range of stimulus material.

The 1972 project

The main aim of this project was to investigate student performance across a fairly wide range of written and visual stimulus material item types.

Table 1: Comparison of candidates' performance in the pre-test and GCE for 1971

		Document paper				Map paper			
		Number	Pre-test Mark	GCE Mark	Correlation	Number	Pre-test Mark	GCE Mark	Correlation
School 1	Mean	20	49·95	53·10	·73	20	59·25	57·75	·71
	SD		11·80	11·50			15·84	13·30	
School 2	Mean	57	60·32	52·38	·64	48	38·68	47·31	·41
	SD		14·05	12·20			11·66	15·09	
School 3	Mean	40	67·45	52·62	·67	43	64·47	51·28	·78
	SD		13·53	13·03			17·08	11·22	
School 4	Mean	26	49·12	50·58	·40	28	53·64	53·43	·32
	SD		10·89	10·44			13·58	10·50	

A secondary aim was to find out what type of stimulus material students chose where a choice was offered. The test was administered to four schools in May 1972. These schools were all defined 'standard' schools by the University of London School Examination Department and included a well established boys' grammar school in central London, a mixed comprehensive in the East Midlands, a girls' grammar school in outer London and a mixed grammar school in the Home Counties. One paper was set comprising twelve stimulus material questions, two on each of six topics as follows:

Questions	*Topic*	*Type*
1	General Strike	Cartoon
2		Documents
3	India	Photograph
4		Map
5	Social Reform	Document
6		Cartoon
7	International rivalries pre-1914	Statistics
8		Map
9	Treaty of Versailles	Document
10		Map
11	Russia	Photographs
12		Document

Thus a wider variety of type of material was employed than in the 1971 project. The topics chosen were again of central importance, three on European, three on English history. In the two hours available students were required to answer three questions, not more than one on any one topic.

Each question comprises a piece of historical evidence followed by six or seven questions. The evidence varied in nature. It included photographs, cartoons, maps, tables of statistics and documentary extracts. Some examples of these are given here.

Cartoon question

Question 1

Study the cartoon (on page 86) about the General Strike and then answer questions(i)–(vi) which follow.

		Marks (max.)
(i)	What office did Baldwin hold in 1926?	1
(ii)	The *Daily Herald;* The *Daily Mail;* The *Workers' Weekly*: which of these publications would have been most likely to publish this cartoon in 1926?	1
(iii)	Explain the reasons for the *Coal Crisis* depicted in the cartoon.	4
(iv)	How would you support the opinion that Mr. Baldwin had 'got us through that fog splendidly'?	5
(v)	Describe the legislation of 1927 with which the Government sought to deal with industrial difficulties.	3
(vi)	Outline the effects of the General Strike during the next five years on (a) trade unionism, and (b) the Labour Party.	6

Document question

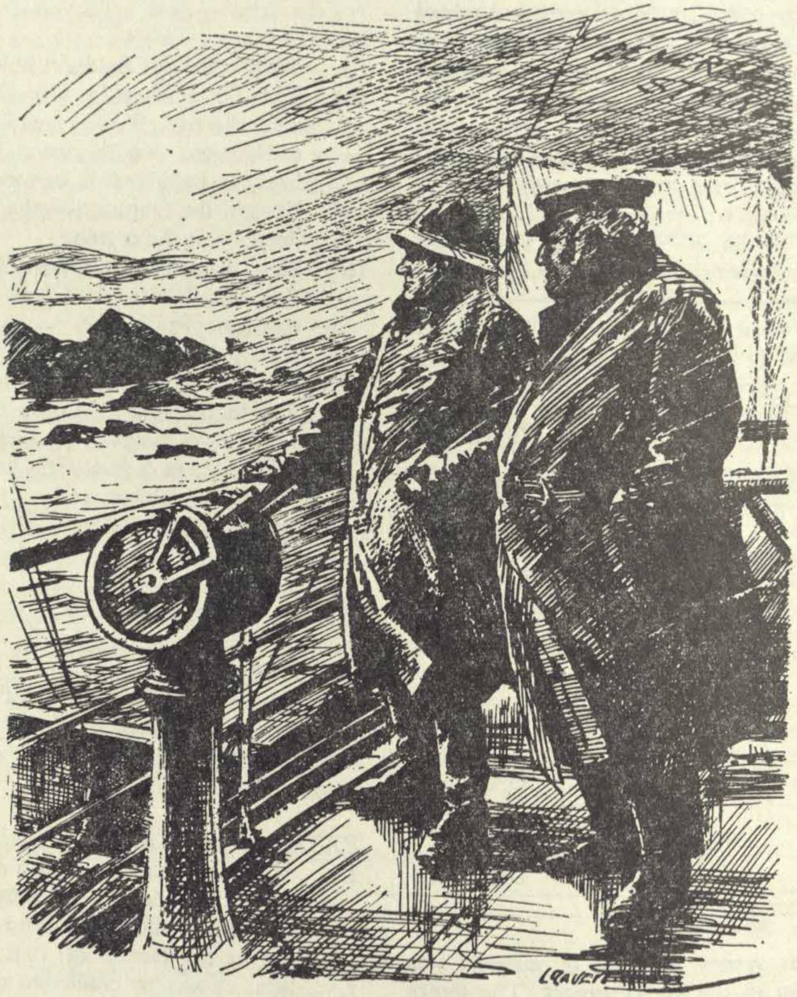
Question 2

Study the following extracts from the *British Gazette* and the *British Worker* at the time of the General Strike and then answer questions (i)–(v) which follow.

EXTRACT A (The *British Gazette*, 6.5.1926)

The strike is intended as a direct hold-up of the nation to ransom. 'This moment,' as the Prime Minister pointed out in the House of Commons, 'has been chosen to challenge the existing Constitution of the country and to substitute the reign of force for that which now exists. I do not believe that there has been anything like a thoroughgoing consultation with the rank and file before this despotic power was out into the hands of a small committee in London. I do not think all the leaders who assented to order a general strike fully realized that they were threatening the basis of ordered government.'

PUNCH, OR THE LONDON CHARIVARI.—MAY 26, 1926.



THE MAN IN CONTROL.

JOHN BULL (*to the Pilot*). "YOU'VE GOT US THROUGH THAT FOG SPLENDIDLY."
MR. BALDWIN (*sticking quietly to his job*). "TELL ME ALL ABOUT THAT WHEN WE'RE PAST THESE ROCKS."

By Permission of *Punch*

EXTRACT B (The *British Worker*, 7.5.1927)

The General Council does not challenge the Constitution. It is not seeking to substitute unconstitutional government. Nor is it desirous of undermining our Parliamentary institutions. The sole aim of the Council is to secure for the miners a decent standard of life. The Council is engaged in an Industrial Dispute. There is no Constitutional Crisis.

The General Council struggled hard for peace. The are anxious that an honourable peace should be secured as soon as possible. They are not fighting the community. They are defending the mine workers against the mine-owners.

Marks (max.)

(i) Which of the above Extracts (A or B) better represented the opinions of Winston Churchill? 1

(ii) What organization was represented by
(a) the *British Gazette*, and
(b) the *British Worker*? 2

(iii) Explain
(a) what the Prime Minister meant by the challenge to *the existing Constitution of the country* (lines 5–6) 2
(b) *a small committee in London* (lines 12–13) 1
(c) how the attempt *to secure for the miners a decent standard of life* (lines 22–23) led to the General Strike 4

(iv) How effective was the General Strike in *defending the mine workers against the mine-owners*? (lines 31–33) 3

(v) Why did the General Strike last for only nine days? 7

Photograph question

Question 3

Study this photograph of Gandhi at a meeting in London in 1931 and then answer questions (i)–(vii) which follow.

[*For technical reasons, it has not been possible to reproduce here the photograph which showed Gandhi, watched by a crowd, as he left the Friends' Meeting House in London after his first meeting in England.*]

Marks (max.)

(i) Explain
(a) why Gandhi came to England in 1931, and
(b) what he achieved. 3

(ii) What significance can be attached to the manner of Gandhi's dress? 2

(iii) Apart from matters of dress, why did Gandhi attract so much attention in England in 1931? 3

(iv) Apart from seeking independence for India, name *two other* causes for which Gandhi campaigned in his life-time and in *each* case indicate briefly what successes he achieved. 4

(v) What weapons did Gandhi use to try to achieve his aims? 2

(vi) In what ways was Gandhi disappointed when Britain granted independence to India? 3

(vii) Briefly explain the manner of Gandhi's death. 3

Statistics question

Question 7

Study the following statistics of selected countries in 1900 and then answer questions (i) to (v) which follow.

	Austria-Hungary	France	Germany	Britain	Italy	Russia
Population (ms.)	45	39	56	42	32	133
Men in regular army	397,000	590,000	585,000	280,000	260,000	860,000
Annual iron and steel production (m. tons)	2·6	3·3	13·8	13·9	0·5	5·0
Annual value of foreign trade (£m.)	152	460	545	878	133	142
Merchant fleet (m. tons)	0·3	1·0	1·9	9·3	0·9	0·6
Battleships	6	23	14	49	14	24

Marks (max.)

(i) State briefly what conclusions you would draw from:
  (a) a comparison of Germany's population statistics with those of Britain and France. In what ways was the character of Austria-Hungary's population different from those of these three?
  (b) a comparison of the relative sizes of the populations and the regular armies in the cases of France, Germany, Britain and Russia
  (c) a comparison in the case of all six countries listed, of the relative sizes of their populations and their annual statistics of iron and steel production. 6

(ii) State the size, according to these figures, of the total regular armies *and* battleship navies, in 1900, of:
  (a) the Triple Alliance powers
  (b) the future Triple Entente powers. 4

(iii) Show briefly what light is thrown by these statistics on Germany's anxiety to:
  (a) avoid a war on two fronts
  (b) step up the rate of her battleship construction. 2

(iv) What evidence do these figures provide that Anglo-German rivalry in 1900 was industrial, maritime and commercial rather than military? 3

(v) Illustrate by reference to *one* overseas crisis, *before* 1914, how Britain and Germany confronted each other as rivals for world power. 5

Map question

Question 8

Study the map (on page 89) of colonial Africa (1900), and then answer questions (i) to (v) which follow.

Marks (max.)

(i) Indicate *by letter* a mainland territory of Africa which in 1900 was
  (a) A British colonial possession
  (b) under British advice and protection, but not a colony
  (c) engaged in a war against Britain. 3

(ii) Indicate by letter a mainland African territory controlled in 1900 by
  (a) France
  (b) Germany
  (c) Italy.

(iii) Name *two* of the following:
  (a) international waterway X
  (b) off-shore island Y, ceded by Germany to Britain in 1890
  (c) off-shore island Z, annexed by France in 1894. 2

(iv) State briefly:
  (a) *three* reasons for which European powers desired colonial possessions in Africa at the end of the nineteenth century
  (b) *three* obstacles that delayed the opening-up of the African continent for so long. 6

(v) Explain briefly the circumstances in which an international crisis developed *either* in area J in 1898, *or* in area A in 1905. How was the crisis you have selected eventually resolved? 6

This stimulus material was intended to trigger off a chain of thought relating to a particular historical topic. Sometimes the answer lay directly within the stimulus material itself; usually this was not so. A final sub-question in each case, invited candidates to go outside the immediate field of knowledge. Thus a wide variety of skills was tested.

Preference

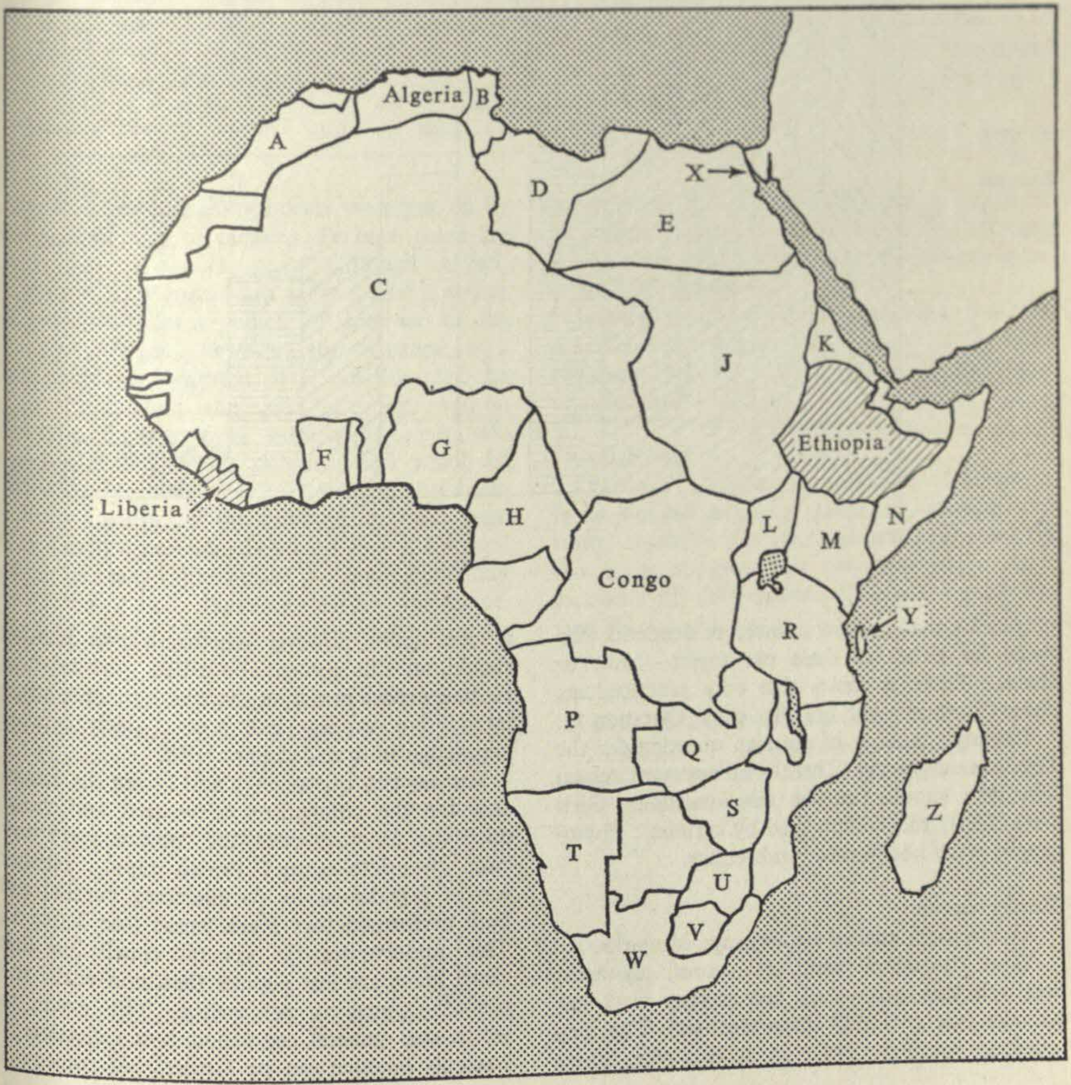
Which types of stimulus material do students prefer? Figure 1 shows the popularity of the questions coded according to the type of material. The statistics question (Question 7) attracted 61 per cent of the candidates. The documents questions (Questions 5 and 2) were ranked second and third with 42·6 per cent and 41·6 respectively. Map question 9 came fourth with 43·2 per cent.

Boys and girls tended to choose differently, Table 2 compares the relative preference for a performance in the twelve questions as between boys and girls.

Table 2 presented a paradox. Whereas in general, preference differed, in both cases the statistics question was more popular. This seemed to contradict an impression common to many history teachers that history students would not be able to do questions based on statistics successfully and that girls in particular would rarely choose this type of question. In fact girls seem to have been put off by the question selected as a pair to this question—a map question.

Documentary extracts were the next most popular questions. But here there is a clear difference in the choice of boys and girls particularly in the mixed schools; the girls preferred the document on Lloyd George's social reforms (this was accompanied by a simple statistical schedule!); the boys showed a strong preference for the newspaper extracts on the General Strike. Possibly this difference is a reflection of male interest in industrial relations and strife, whilst women are more favourably disposed towards the constructive aspects of social reform within a legal framework.

Generally, visual material questions (maps, cartoon, photographs) were markedly less popular than non-visual ones. In retrospect, however, at least two of the visual items (the photograph items: Question 3 and 11) were on topics which are apparently peripheral to the average syllabus



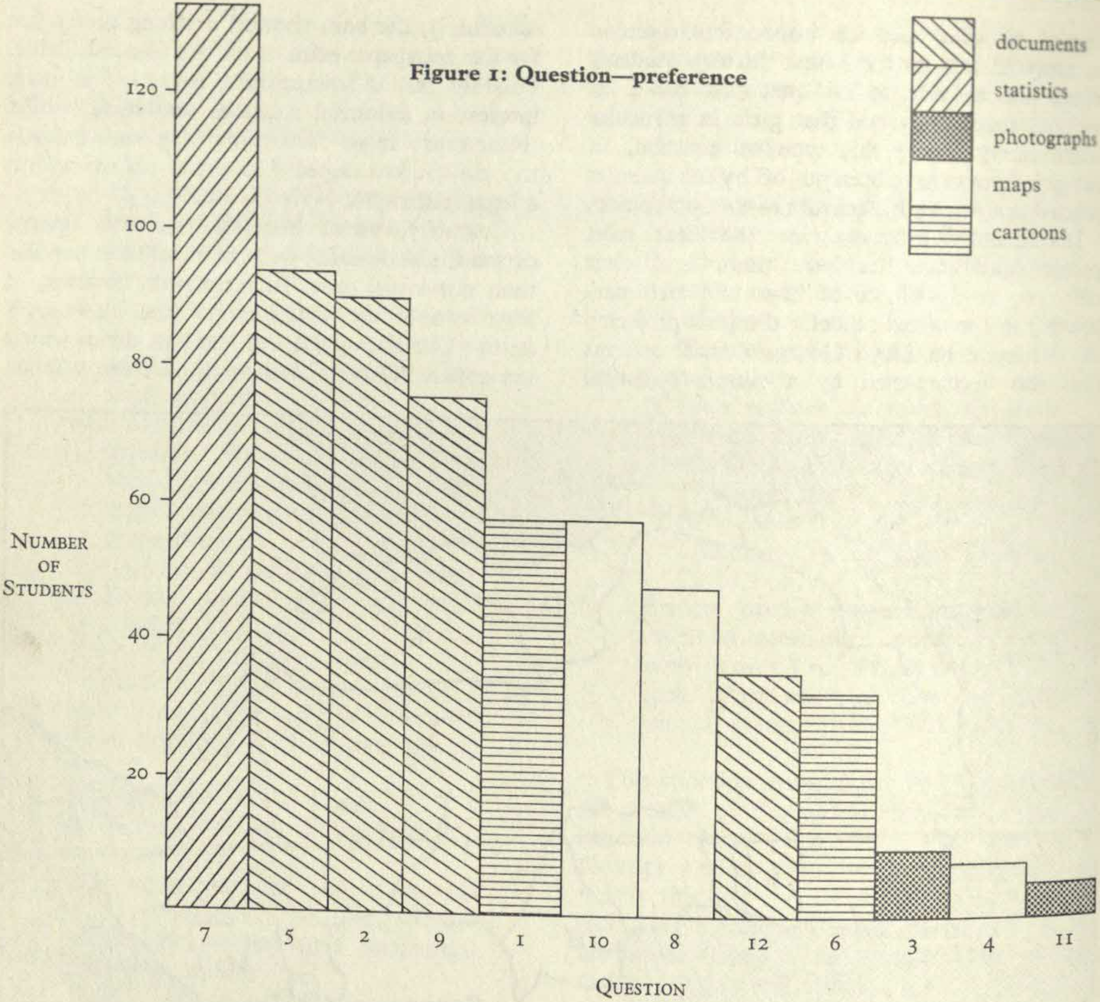


Figure 1: Question—preference

judging by the very low number of students who opted for either of these two topics—India or Russia. Table 2 shows that only ten students chose Question 3 and only five chose Question 11.

The high ranking of the map question for the boys was surprising. There is no apparent reason why girls should not like map questions. Girls were clearly more stimulated by cartoons. Photographs were not popular with either.

Performance

All questions earned a wide range of marks.

Visual questions were, *in general*, answered less well than non-visual, just as they were less popular (the two tend to go together). The map questions were the exception to this rule. There is some doubt, however, whether a map question should, in this context, be called a visual item for its purpose is not on the whole to start off an imaginative response but to elicit fairly detailed knowledge in a well-defined field.

The cartoon questions (1 and 6) are the crucial ones for the comparison of written and visual questions. In both cases, the subjects were very basic twentieth-century British topics. Neither was as popular as its documentary parallel and both performed very much worse (see Table 2). This underwrites the general conclusion that visual items are less popular than non-visual and perform poorly in comparison. The extent to which this general conclusion is a reflection of teaching practice can be a matter of conjecture

Table 2: Preference and performance of the 12 questions

QUESTION	GIRLS			BOYS			TOTAL			TYPE
	N	Mean	SD	N	Mean	SD	N	Mean	SD	
1	48	6·94	3·62	10	10·90	3·94	58	7·62	3·97	Cartoon
2	55	9·00	3·41	36	11·00	3·74	91	9·79	3·68	Document
3	0	0·0	0·0	10	8·80	3·60	10	8·80	3·60	Photograph
4	1	3·00	0·0	8	10·25	5·04	9	9·44	5·27	Map
5	67	10·21	2·72	26	8·81	3·71	93	9·82	3·77	Document
6	25	7·16	3·09	8	7·00	1·94	33	7·12	2·86	Cartoon
7	76	11·88	3·67	57	11·77	4·13	133	11·83	3·87	Statistics
8	18	11·72	4·16	30	12·07	3·93	48	11·94	4·02	Map
9	46	9·98	3·29	29	9·79	3·13	75	9·91	3·23	Document
10	26	10·35	3·16	32	10·75	4·01	58	10·57	3·66	Map
11	2	9·00	0·0	3	7·33	3·86	5	8·00	3·10	Photograph
12	13	10·85	4·33	23	12·61	4·42	36	11·97	4·47	Document
Total	126	9·80	3·94	92	10·78	4·12	218	10·21	4·04	

only. It is possible that students were put off by the unfamiliarity of cartoons. Perhaps these are rarely used in the classroom, although in fact, they are a very fruitful and easily available source of evidence offering much of interest to the student. Perhaps, however, the situation is a little more sophisticated. It is possible that the more able student who could have done well on this type of question deemed it wiser to play safe and go for the sort of question with which he was slightly more familiar. Weaker brethren, not having studied the topics fully, may have been fatally lured into believing that a 'picture' offers an easy alternative. A third possibility could be called the examination syndrome. Students near to the GCE examination may be conditioned towards a totally literary approach. It is to be hoped that this is not so. In any case, such a situation will change with the change in the style of examination.

Interestingly, the statistics question was not only the most popular but also one of the most successful questions with both girls and boys scoring equally highly.

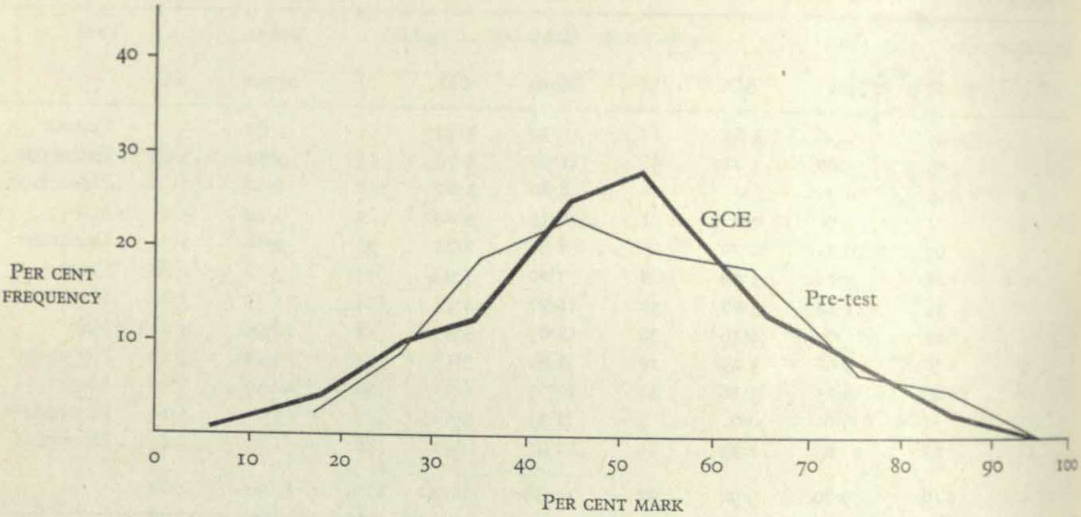
Question choice

Choice of question and question combination is an indication of the effect of teaching and study methods. Distinct patterns of question choice are apparent. *Six combinations—out of a maximum of 160—account for one third of the total number of students* (see Table 3).

Table 3: The most popular question combinations

Question Combination	No. of Candidates	Average Pre-test Mark %	Average GCE mark	Correlation Pre-test/GCE Mark	Question types
2, 5, 7	22	49·08	47·32	·742	D D S
1, 5, 7	17	46·47	51·24	·729	C D S
2, 7, 9	14	52·18	47·57	·211	D S D
5, 7, 9	12	56·38	53·17	·624	D S D
2, 8, 10	9	66·12	46·17	·012	D M M
2, 7, 12	8	56·25	49·25	·448	D S D
Total sample	219	50·73	49·26		

Figure 2: Performance distribution of pre-test and GCE marks



The most common question combination (2:5:7)

This combination was entirely non-visual and it is interesting that the performance of these candidates in the pre-test compares well with their GCE mark, the correlation being ·74. On the other hand there is no such relationship for the group offering the highest scoring combination (2:8:10) comprising 2 maps and a document. All the students selecting this combination were from the same school and all did very well, particularly on account of their apparent familiarity with map work. Interestingly those candidates choosing the same topics but different questions (2:7:9) also show little relationship with their GCE performance.

The performance of the group containing the cartoon question (1:5:7), although second in popularity was the lowest scoring. This group, all girls, clearly found the cartoon very difficult and suffered for this in their overall score. Nevertheless the performance of this group correlates well with their GCE mark.

It is evident from Table 3 that GCE performance can not always be predicted from a candidates' pre-test score. The performance will vary and it will depend, to some extent on the choice of question. The response pattern for the stimulus question differs from the traditional essay type. Significantly, Question 7 appears in five out of the six most popular combinations.

Comparison of pre-test with GCE

The two types of test seem to perform differently. While the average marks for the two tests are much the same, 50·73 per cent for the pre-test and 49·26 per cent for GCE, the pre-test distribution, Figure 2, has a wider spread of marks and hence tends to give better discrimination between candidates.

Boys and girls perform differently on the two types of paper. Table 4 shows that the boys perform better on the structured stimulus questions than on the essay questions of the 'O' level examination. The position is the opposite for the girls.

Table 5 shows the correlation between the

Table 4: Pre-test and GCE marks

	Boys			Girls			Total		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Pre-test	92	50·31	17·80	127	47·33	15·11	219	50·73	16·43
GCE	92	47·85	15·48	127	50·28	16·55	219	49·26	16·15

Table 5: Correlation between pre-test mark and GCE mark by school

School	Type of School	Correlation	N
I	Boys	·59	43
II	Mixed	·52	66
III	Girls	·69	63
IV	Mixed	·41	48
Total sample		·53	220

two marks for the four centres and the total sample.

There is a significant relationship between the two tests.

If we divide the candidates sitting the pre-test into ability bands on the basis of their performance in the pre-test and the operational examination, several differences emerge.

In the pre-test more boys fell into the highest ability band and fewer into the lowest ability band than in the GCE examination. The reverse is true of girls; there are 22 girls in the highest ability band in the GCE examination compared with only 9 in the pre-test. Why is it then that boys do seem to do better than girls at this type of test?

If we consider the 23 boys who did Question 12 and look at the performance in this question by ability level, most of them fall in the lower GCE ability ranges. Yet, by doing this question, they are able to raise their marks on the pre-test. Even the people in the lowest ability band, as measured by their GCE performance, are able to achieve a mean mark of 9·1 out of the possible 20 marks allotted.

We have seen that the question combinations (1:5:7), (2:5:7), (5:7:9) and (2:7:9) are the most popular, as one third of all the candidates sitting the pre-test chose these. Comparing the GCE mark with the pre-test mark for these four groups shows that the brightest candidates are likely to choose Question 5, 7 and 9. Looking at these questions in detail confirms this—a large percentage, 68 per cent, of the highest ability band

Table 6: Number of candidates in each ability band for the pre-test and GCE examination

		PRE-TEST		
		Boys	Girls	Total
Ability Band 1	lowest	15	26	41
2		21	29	50
3	average	18	35	53
4		19	28	47
5	highest	19	9	28
Total		92	127	219

		GCE		
		Boys	Girls	Total
Ability Band 1	lowest	20	24	44
2		21	21	42
3	average	16	31	47
4		25	29	54
5	highest	10	22	32
Total		92	127	219

chose Question 5 and 50 per cent chose the other two questions.

Points to note

1. Stimulus questions discriminate vigorously giving a wider spread of marks than the conventional essay questions.
2. Boys and girls appeared to perform differently in the pre-test. Stimulus questions with short-structured answers seem to favour the boys.
3. The candidates showed a preference for the verbal stimulus. Perhaps they preferred to play safe with a document question rather than experiment with an unfamiliar cartoon or map and of course there has been a long period of verbal habituation in the examining of history.
4. The various types of stimulus question performed differently. In general performance was best on the verbal type of question. Against expectation, candidates were not afraid of the statistics question and did well on it but map questions were not popular, particularly among girls.
5. The three experiments carried out show that the stimulus question can be used profitably in history examining. Candidates seem to be able to understand the questions and answer them well.
6. More research is being carried out to investigate the use of stimulus material questions at 'A' level.

Adding and Averaging Grades

John Nisbet, Department of Education, University of Aberdeen

Summary

Letter grades in marking have to be converted to a numerical scale before they can be added or averaged, and the conversion may introduce error in the final total. The use of the stanine scale to control some (but not all) of these sources of error, is described, together with a simple graphical conversion procedure.

Introduction to the problem

Letter grades (A, B, C, etc.) are often used in marking essays or examination answers. Many teachers who are properly sceptical of marking in percentages because this gives a false impression of precise discriminations, prefer to use a coarse scale of five or six letter grades. Grades are often sufficient as an indication of level of performance, and it may seem easier to mark consistently on a coarse scale like this than on a scale from 0 to 20, or from 20 to 90 per cent. However, a difficulty arises when several such grades have to be combined to give an over-all mark, as when assessments over a whole session are averaged to give a mark for the year, or when marks for separate examination answers are added to give a total for the whole paper.

The usual method of combining grades is a crude system of equivalent numbers. Letter grades are converted to numbers (A=4, B=3, C=2, etc.) which are then added or averaged. This procedure may give false results, for it assumes that the intervals between grades are equal units. That is, it assumes that the difference in standard between A and B is the same as the differences between B and C, C and D, and so on, since each of these differences counts as one unit in the total. There are occasions when this assumption is clearly invalid. In the scale illustrated in Table 1, for example, the grades do not represent equal intervals. The four grades, D, E, F and FF are close together, while at the other end of the scale the four grades A, B, C and D (or even A, B, C+ and C) cover a very wide range of performance. Also, if different markers interpret the meaning of letter grades differently, this, too, will introduce error when grades are combined. On an A–E scale, one teacher may use C as the minimum pass standard, while others work to some version of a normal distribution, which may be 10, 25, 40, 25, 10 per cent respectively, or 5, 20, 50, 20, 5. (A similar criticism can in fact be applied to percentage marks, though the fallacy is concealed because it seems that we are dealing with absolute numbers.)

Table 1: Example of a marking scale with unequal intervals

Grade	*Label*	*Percentage of class given this grade*
A	Distinction	approx. 5 per cent
B	Merit	15
C+ C C–	Satisfactory	60
D	Doubtful pass	5
E	Borderline	5
F FF	Fail Outright fail	No fixed proportion

In order to convert letter grades into numbers which can be added or averaged, it is necessary to match the grades against some scale of 'standard scores'. An explanation of the principle of 'standard scores' can be found in Guilford (1956). Standard scores are based on the percentage distribution of a set of marks or grades, and the equivalences are calculated so that the resulting standard scores are in a normal distribution where unit intervals may validly be assumed to represent equal intervals in standard of performance. There are various forms of standard score, and the one recommended here is a stanine scale in which there are nine categories. The stanine scale was devised for use by the US Army Air Force Aviation Psychology Program during World War II, and the percentage distribution of stanines is set out in Table 2.

Table 2: The stanine scale

Stanine	9	8	7	6	5	4	3	2	1
% of total in each stanine	4	7	12	17	20	17	12	7	4
Cumulative %	100	96	89	77	60	40	23	11	4
Percentile midpoint	98	93	83	69	50	31	17	7	2

Though the stanine scale has predetermined proportions at each level, there is no constraint on the actual marking. Teachers may use whatever marking system and whatever distributions they find appropriate; the grades they give are then converted into stanines for the purpose of adding or averaging—and the result can be converted back to the grades which are meaningful to the teachers. This is a well-established practice in educational research when teachers' marks are combined to give a criterion of school performance. Teachers themselves may hesitate to apply it because the computation seems time-consuming. It is in fact relatively straightforward, and this article outlines a graphical method which some may find simpler.

The conversion procedure removes the idiosyncracies of different grading practices so that no one set of marks has undue influence in the total. The procedure cannot affect the order of merit within any one set of marks, and it will not produce dramatic changes when marks are added: a pupil or student with low marks or grades in each question will end up with a low total mark. But at the pass–fail level, or wherever a cut-off is imposed, it is important that a pupil's fortunes should not depend on the chance operation of oddities in the marking. For example, if a teacher gives most of his marks or grades over a narrow range, then one very low grade (such as zero for misunderstanding a question) can pull a good candidate down so much that he cannot make up the lost ground, or a very high mark in one single question can give someone an unbeatable lead. A teacher who uses marks over a wide range, or a question which spreads the candidates, has more weight in the total than when marks cover only a limited range. The stanine conversion corrects for effects such as these. It may even over-correct: perhaps one pupil deserves an unbeatable lead, or perhaps all the marks on one answer should be all about the same—the stanine scale treats each set of assessments as comparable.

Procedure

In this example, three sets of grades are to be added to give a total. The frequency distributions are:

	Assessment in:		
Grade	First term	Second term	Third term
A	6	5	11
B	15	6	19
C	36	40	30
D	9	11	9
E	4	5	6
F	2	7	—
FF	3	1	—
Total	75	75	75

Step 1. Count the frequency distribution for each set of grades (as above).

Step 2. Calculate a cumulative frequency, and convert to a cumulative percentage frequency. For example, using the first term assessments above:

Grade	*f*	*Cf*	*Cf%*
A	6	75	100
B	15	69	92
C	36	54	72
D	9	18	24
E	4	9	12
F	2	5	7
FF	3	3	4
	75		

Step 3. Express the cumulative percentage as a range for each grade, and calculate the midpoint of each range. (Short cut method: Add 0 + 4, divide by 2; add 4 + 7, divide by 2; add 7 + 12, divide by 2; etc.)

Grade	*f*	*Cf*	*Cf%*	*Range*	*Midpoint*
A	6	75	100	92–100	96
B	15	69	92	72–92	82
C	36	54	72	24–72	48
D	9	18	24	12–24	18
E	4	9	12	7–12	9·5
F	2	5	7	4–7	5·5
FF	3	3	4	0–4	2
	75				

Step 4. Opposite each midpoint write the nearest stanine midpoint, which is got by reference to Table 1 above (bottom row).

Grade	*f*	*Cf*	*Cf%*	*Midpoint*	*Nearest stanine midpoint*	*Equivalent stanine*
A	6	75	100	96	98	9
B	15	69	92	82	83	7
C	36	54	72	48	50	5
D	9	18	24	18	17	3
E	4	9	12	9·3	7	2 (or 2½)
F	2	5	7	5·5	7	2
FF	3	3	4	2	2	1
	75					

The stanines in the extreme right column are the equivalents of the grades in the extreme left column. These stanines can be validly added to any other stanines (for this class) and averaged.

(*Note:* E and F grades are so close that even on a nine-point scale the difference does not amount to one unit of measurement. They can either be given the same stanine equivalent, or a half can be used.)

Alternative graphical procedure

The use of a prepared graph paper simplifies computation in Steps 3 and 4 above. The vertical axis in the graph is cumulative percentage frequency. The horizontal axis is the scale of marks or grades being used. Horizontal lines at 4, 11, 23, 40, 60, 77, 89 and 96 per cent mark off the stanines (see Figures 1, 2, 3 as examples).

To illustrate from the set of marks used in the previous example:

Step 3. On the graph, for column FF, draw a line at 4%. Block in 0 to 4%. For F column, draw a line at 7%. Block in 4 to 7%. For E column, draw a line at 12%. Block in 7 to 12%, etc. Place a dot in the middle of each block. Note, from the horizontal lines and the stanine numbers on the right-hand side, which stanine the dot falls into. This is the stanine equivalent for that grade (see Figure 1).

As further illustrations, the set of second term assessments in the example is converted to stanines by the computational method, and the third term assessments are dealt with by the graphical method (Figure 2).

SECOND TERM ASSESSMENTS

Grade	*f*	*Cf*	*Cf%*	*Midpoint*	*Nearest stanine midpoint*	*Equivalent stanine*
A	5	75	100	96·5	98	9
B	6	70	93	89	93	8
C	40	64	85	58·5	50	5
D	11	24	32	24·5	31	4
E	5	13	17	14	17	3
F	7	8	11	6	7	2
FF	1	1	1	1	2	1
	75					

THIRD TERM ASSESSMENTS

Grade	*f*	*Cf*	*Cf%*	*Equivalent stanine*
A	11	75	100	8
B	19	64	85	6
C	30	45	60	4·5
D	9	15	20	3
E	6	6	8	1·5

In the case of the third term assessments, because the frequency distribution approximates to a normal distribution, the use of a crude number equivalence (A=8, B=6, C=4, D=2, E=0) would give roughly comparable figures at the

Figure 1

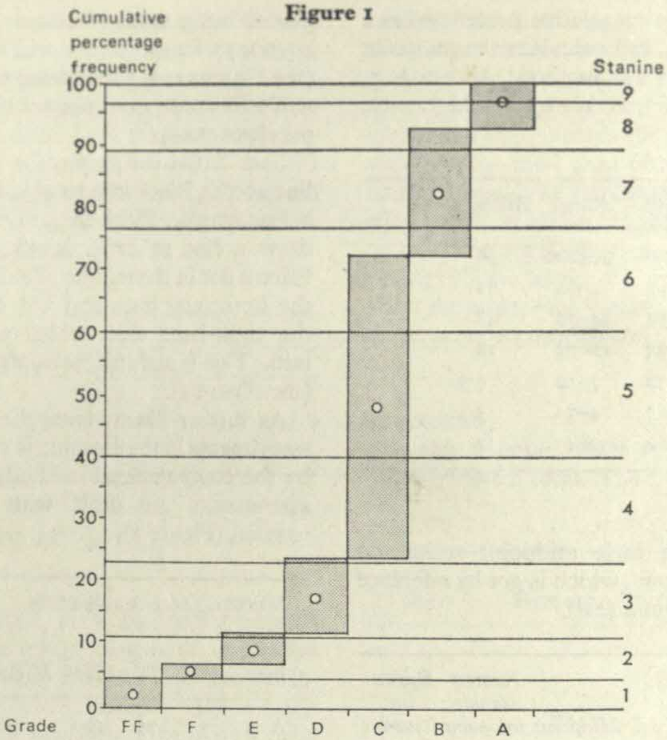
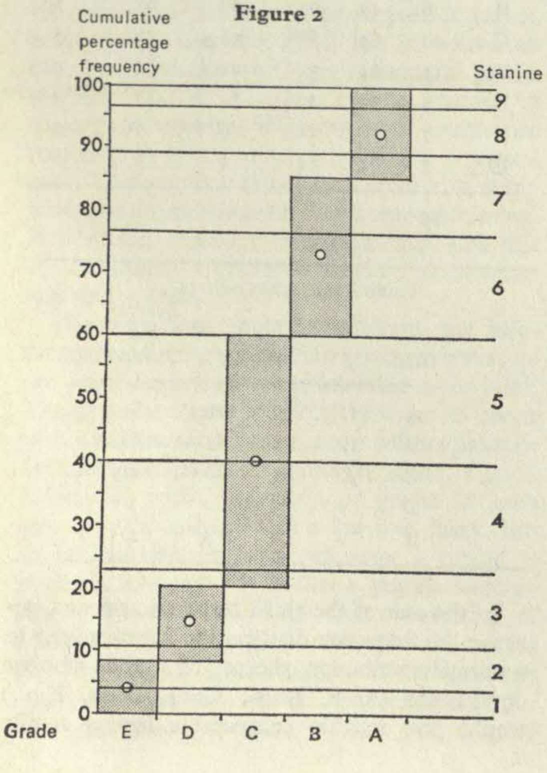


Figure 2



upper end of the scale, unlike the other two sets of assessments in this illustration.

Application to numerical marks

Even when numbers are used in marking, the straight addition and averaging of marks may be invalid. This procedure can also be used for smoothing out inequalities in sets of numerical marks. For example, an examination requires answers to five questions from a choice of eight. Different teachers mark each question out of 20 marks, and the totals are added to give a 'percentage' mark out of 100. Some markers may be severe, others lenient; some may use the whole range from 0 to 20, others marking only between 5 and 15. Conversion of each teacher's marks to a standard scale removes the differences between markers. (It may do this too well, removing differences which ought to remain—if, for example, the weaker students all do Question 3 and only the most able students do Question 5. It is possible to make an allowance for this, adding or subtracting one stanine to all marks for certain questions; but these adjustments go beyond the scope of this paper.)

The following example shows the conversion of one set of marks for 160 pupils on a question marked out of 20. The procedure can be applied to grouped scores (0–4, 5–9, 10–14, etc.); in this case, the upper boundary of each interval is half a point above the highest mark in the interval (4·5, 9·5, 14·5, etc.).

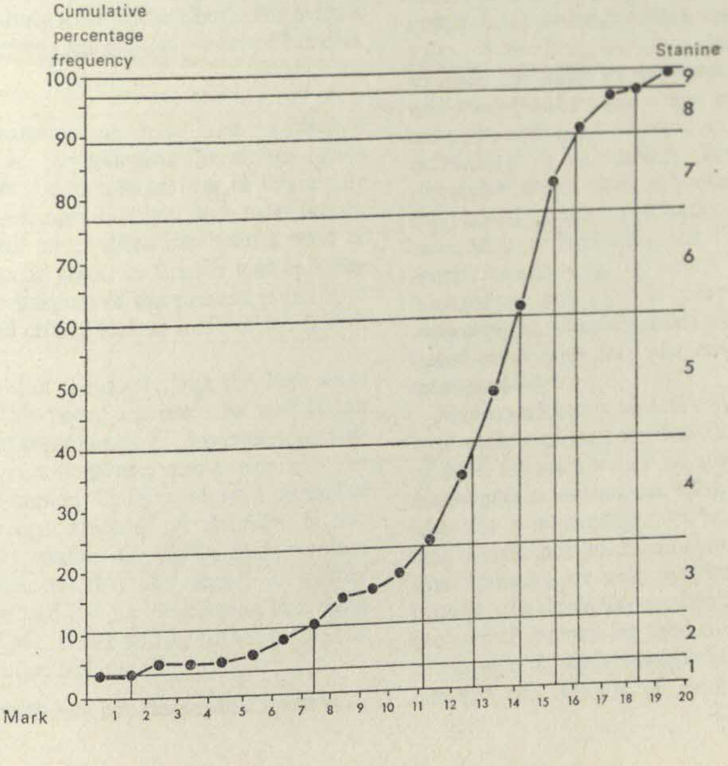
Steps 1, 2, 3. Write down frequency distribution of marks. Express as cumulative frequency. Convert to percentage cumulative frequency.

Mark	*f*	*Cf*	*Cf%*
20	1	160	100
19	3	159	99
18	1	156	97
17	10	155	96
16	15	145	90
15	30	130	81
14	24	100	62
13	20	76	47
12	16	56	34
11	10	40	24
10	3	30	18
9	2	27	16
8	7	25	15
7	4	18	11
6	4	14	9
5	2	10	6
4	0	8	5
3	0	8	5
2	2	8	5
1	0	6	4
0	6	6	4
	160		

continued

Step 4. Plot the percentage cumulative frequency on a graph, with the vertical axis for the cumulative percentage frequency and the horizontal axis for marks 0 to 20 (see Figure 3). Note that the upper boundary of each mark in the cumulative frequency is the mid-point above that mark (a mark of 4

Figure 3



extends up to 4·5). The horizontal lines for stanines 1 to 9 cut the cumulative frequency curve into nine segments. For each segment, read off (by dropping perpendicularly to the numbers at the foot of the graph) the marks which are equivalent to each stanine.

Mark	*Equivalent stanine*
20–19	9
18–17	8
16	7
15	6
13–14	5
12	4
11–8	3
7–2	2
1–0	1

If it is necessary to fix a pass mark or other gradings, this must be done as a separate exercise. Stanines have an average of 5·0. In any one set of stanines, a pass mark of 3 will place approximately 11 per cent below pass (see Table 2). But when several sets are averaged, there is a regression towards the mean, with relatively fewer individuals appearing in the extreme high and low stanines. Thus, less than 11 per cent will fail when an average of 3 is the pass. A pass mark of 4 will probably fail slightly less than 20 per cent. Deciding a minimum pass standard properly involves scrutiny of marginal papers at stanine 2 or 3 or whatever is appropriate as a first approximation and deciding which stanine is above or below the minimum acceptable.

The explanation of these procedures in words is tedious, but the actual working is simple. When assessment is of a general kind, as a feedback to the pupil or student, there is no need to apply procedures of this kind. But when an assessment is the basis for an important decision, this relatively straightforward procedure will at least control one of the sources of random error.

Reference

GUILFORD, J. P. (1956), *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill, p. 503.

Attitude Measurement

a Critique of Some Recent Research

P. L. Gardner, Faculty of Education, Monash University, Victoria, Australia

Summary

This article examines some recent British research in the field of attitude measurement, and contends that a few of these studies display serious faults in instrument design and usage. Effective research in this field demands the construction of attitude scales which clearly reflect some underlying theoretical construct; it requires distinct variables to be kept separate and not added together to produce meaningless total scores; and it requires instruments chosen as outcome measures in research on treatment effects to display some defensible connection with the treatment being studied.

This article was written during 1973 while the writer was on study leave at the Centre for Science Education, Chelsea College, University of London.

Introduction

Attitude measurement has been a central concern of social psychologists for almost half a century, and its adoption as an area of interest by educational researchers is only a little more recent. The field is still alive: there has been a wave of research in recent years, in Britain and in other countries. The purpose of the present paper is to take a close and critical look at some of the British research.

Let it be said at the outset that the arguments presented in this paper are not new and reflect little more than commonsense. However commonsense, to be commonsense, must not only be sensible, but common. Unfortunately, a perusal of some of the recent research on attitudes in this country reveals studies displaying such startling technical deficiencies that one begins to wonder whether the procedures for developing and using attitude scales *are* in fact widely understood. The defects can be classified into three types:

(1) scales which lack any discernible underlying theoretical construct;

(2) scales in which various theoretical constructs are confounded together, i.e. scales which attempt to reduce multi-dimensional attributes to single scores;

(3) experimental treatments in which there is little discernible relationship between the experimental treatment applied and the scale used to measure its outcomes.

The first two defects apparently spring from a lack of appreciation of what measurement *means*. When we measure a physical property of some object—let us take as an example a dining-room table—we assign numbers to some attribute of the object in accordance with some set of operating rules. Note that, strictly speaking, we do not measure *the table*: we measure some well-defined and conceptually clear aspect of it. We may, for example, measure its length, or its weight, or the reflectivity of its surface. If a bright young researcher comes along and tells us that he thinks the rate of tarnishing of cutlery is related to the tarnability of table-tops, the onus is on him to specify exactly what he means by the concept of 'tarnability' and how he proposes to go about measuring it. Clearly defined concepts and operational procedures are indispensable to the act of measurement.[1]

Measurements on objects can be operated upon by a variety of mathematical rules. The outcomes of such operations may be *useful* or *useless*, and *meaningful* or *meaningless*. Obviously, it is impossible for a measurement to be both meaningless and useful, but other combinations are possible. For example, we may multiply the length and breadth of a dining-room table in order to calculate the area of laminate required to resurface it. Such a calculation is both meaningful and useful: the concept of 'area' is of proven value both in science and in the world of practical affairs. In contrast, consider the number obtained by adding together the weight and the reflectivity of the table. Such an operation yields absolutely no information; the number is uninterpretable, meaningless, and hence useless.

Some combinations however yield meaningful concepts which are nevertheless useless. Let us, for example, invent the concept *Biological Size Index*, defined as the length of an animal multiplied by its weight. Such a concept is clearly meaningful: here, combined together into a single number, is an index that describes how 'big' an animal is. Moreover, the concept can be reliably measured. And it is valid: on a BSI scale, elephants would be high, mice would be low, and educational researchers somewhere in between. However, the fact that no biologist has ever proposed measuring animals' BSI scores indicates that there is no apparent pay-off in performing such an operation. In short, measurements should not be combined together unless there are compelling theoretical or practical reasons for doing so.

Two principles, perhaps banal but nonetheless important, have so far emerged from this discussion: measurements should reflect clear concepts, and measurements should not be combined together without reason. It is time to examine in detail some recent studies in which these principles have been ignored.[2]

Defect 1: Lack of theoretical construct

An instrument developed by Harrison (1971) exemplifies this kind of defect. The instrument sought to assess the attitudes of students in a college of technology to their course. Now, it is clearly possible to define theoretical constructs such as 'satisfaction with course' and develop suitable measuring instruments for such constructs: see, for example, the work of Genn (1970), who investigated the satisfaction of university chemistry students with their course.

The problem with Harrison's instrument is that the items do not appear to reflect any clear underlying construct. Consider some of the items:

(a) 'Travelling to this college is a difficult business.'

(b) 'No fees should be charged for my course here.'

(c) 'My lecturers tell me what to learn and no more than that.'

(d) 'I am keen to make a career for myself in teaching.'

Now, in what way does agreement or disagreement with these statements reflect one's attitudes to a college course? Does agreement with (a) reflect satisfaction ('It's hard to get here, but worth all the effort') or dissatisfaction? Does agreement with (b) reflect attitudes to college, or one's economic condition? Does agreement with (c) reflect satisfaction ('I'm a dependent type and I like to be told what to do') or dissatisfaction ('The course is far too narrow')? In a college which is training students for a variety of occupations, how does agreement with (d) reflect one's attitude to the college? The items quoted are not isolated examples: other items in the instrument refer to punctuality, pleasing one's parents, college sport, duplicated notes and so on.

Items such as these might more properly be found in an opinion survey, in which the researcher tabulates and perhaps intercorrelates (but certainly does not sum) the separate responses. They are not the stuff of which attitude scales are made. The absence of any theoretical rationale makes attitude scale theory entirely inapplicable, and statistical procedures such as summing of item scores, split-half reliability[3] and so on are completely irrelevant.

Exactly the same criticism applies to a scale entitled *Attitude to Science Course at Training College* (Shah, 1962). Consider the following items from the scale:

(a) 'I feel that practical work in science should be done in groups.'

(b) 'I feel that the science course at college is too easy.'

(c) 'I feel that I would have preferred to do three separate courses, viz. physics, chemistry and biology rather than an integrated course.'

Again, these items deal with quite separate issues: in this case group work vs. (presumably) individual work, ease/difficulty and specialization/integration.

In both of these studies, there is no single construct underlying the scale. In each case, an extremely high score would indicate a good match between the respondent's opinions and the scale constructor's model of what constitutes a 'good' attitude. Similarly, an extremely low score would point to a complete mismatch. However, an intermediate score would be quite uninterpretable. No purpose is served by applying summated rating techniques in such cases: before one can add up the responses to separate items to produce a meaningful total score, one needs to ensure that the items do have something in common. To reiterate: the act of measurement demands preliminary conceptualizing about what is to be measured.[4]

Defect 2: Confusion of variables

This defect is less elementary than the first, in that there has generally been some attempt to ensure that the items possess something in common. A case in point is the scale constructed by Selmes (1971) and used by Wilmut (1971 and 1973). Some of the items in the scale are:

(a) 'The cold dispassionate scientist is a mythical animal.'

(b) 'Science is a fixed and clearly defined body of knowledge.'

(c) 'There is no such thing as unprejudiced observation; every act of observation we make is biased.'

(d) 'Science is boring.'

The scale purports to measure 'attitude towards science and scientists'. This assumes that the attitude is a single, unidimensional trait. The thrust of the argument which follows is that this assumption is unwarranted and false, and that the single total score yielded by the scale represents a confused mixture of separate variables. In other words, it is a kind of attitudinal *Biological Size Index*. To sustain this argument, we need to examine the rationale of the scale in very close detail.

Wilmut (1971) lists twenty components of scientific attitudes identified by Diederich (1967); scepticism, faith in ability to solve problems, desire for experimental verification, precision, humility, aversion to superstition, and so on. Such a list merely serves to identify *attributes* that scientists have, or are presumed to have. They are logically distinct attributes (just as length, weight and reflectivity are logically distinct attributes of dining-room tables). They are not necessarily unrelated attributes: people's scepticism, for example, might be correlated with their aversion to superstition (just as longer tables tend to be heavier tables). However, many of the attributes listed are likely to be psychologically distinct: humility and desire for experimental verification, for example, are probably unrelated (just as the length and reflectivity of tables are unrelated).

If this argument is accepted, then it is conceptually muddy in the extreme to pretend that there is a single, unidimensional trait called 'the scientific attitude' which can somehow be measured by adding up scepticism plus faith plus precision plus humility. . . Yet this is precisely the argument that Wilmut puts forward. Following his presentation of Diederich's list, he asserts:

> 'Thus, a scale which seeks to determine the attitudes of sixth-form students towards science and scientists would need to contain items which are relevant to each of these attributes.' (1971, p. 82.)

In short, it seems that the way to measure a multi-dimensional entity is to list its separate attributes and then add them all together. This is a very curious argument indeed, akin to arguing that the best way to quantify the properties of dining-room tables is to add up their length, weight, reflectivity . . . and so on.

Wilmut does not actually develop a scale based on Diederich's list. Instead, he refers to a shorter list of attitudes identified by Aiken and Aiken (1969) in their review of the literature on attitudes to science. For brevity, these attitudes can be summarized as: (a) attitudes towards scientists, (b) attitudes towards science, and (c) understanding of the scientific method.[5] Wilmut accepts this list as the basis for the scale he employs.[6] His procedure now is to adopt a scale already devised by Selmes (1971). This scale contains items of four types:

(a) items which refer to scientists and their work;

(b) items which refer to science in general terms;

(c) items which state an opinion about science as a method of investigation;

(d) items which are either emotive in form or thought to have a latent relationship with other items.

(The four items quoted earlier illustrate these four item types.)

The first three types of items correspond to the categories in Aiken and Aiken's review; the theoretical construct underlying the fourth is a little difficult to discern.

The feature of Selmes' scale which makes it so technically defective is that all the items are treated as if they were contributing to a unidimensional trait, i.e. a student's 'attitude to science and scientists' is the sum of his responses to all the items in the instrument. Here we see in action the argument put forward by Wilmut following his discussion of Diederich's list, namely that the way to measure attitudes is to identify attributes and then add them all together. The implicit assumption is that attitudes to science, attitudes to scientists and understanding scientific method are all just aspects of a single variable.

Surprisingly, Selmes actually considers that these attitudes might be separate variables, and then dismisses the possibility, or at least does nothing about it:

> 'With four types of items, it might have been possible to form sub-scales based on each type of item: however, this approach was not adopted for two reasons—this method of item analysis gives no basis for such an approach (only factor analysis would provide the kind of correlational information which is needed) and many items appeared to differentiate between post-graduates but not sixth-formers, and vice versa' (p. 74).

The method of item analysis employed was to calculate Edwards' *t*-scores for each item, high and low groups being identified on the basis of total scores on all items. Selmes is quite correct when he asserts that such an analysis cannot be used to allocate items to sub-scales, but his assertion that factor analysis is the *only* way to perform such an allocation is quite incorrect. His second 'reason' is not a reason at all: whether items on a scale discriminate differently for different age groups is unrelated to the issue of whether different variables should be added together.[7]

There are at least two techniques, apart from factor analysis, that might have been adopted to allocate items to sub-scales. One involves a *sorting* procedure, the other a *rating* procedure. Both require the formation of a panel of competent judges, and both require the researcher to specify, in advance and as clearly as possible, descriptions of the theoretical constructs he is trying to measure with his scales. In the first technique, judges are independently presented with a set of descriptions of the attitude scale constructs and a jumbled collection of items, and are asked to identify the construct appropriate to each item. Items on which judges display widespread disagreement can be eliminated. Gardner (1972) has used this sorting procedure in classifying items describing teacher behaviours into eight different categories.

In the second technique, judges are presented with a single scale description, together with all the items supposedly belonging on that scale. Thurstone-type procedures are then applied (even though the scales might ultimately be scored using Likert-type procedures.) The judges rate each item, often on an 11-point scale, on a continuum running from an extremely favourable attitude to an extremely unfavourable attitude. (The rating refers to the strength of the attitude expressed by the item, and not the judges' personal opinions.) When this procedure is carried out, ambiguous or irrelevant items tend to be given either a wide range of ratings or a fairly neutral median rating. Such items can be eliminated. The same procedure is then used for each of the other scales. Items which survive can then be randomly assorted within a single instrument; the separate scales are of course scored separately. Gardner (1972) has used this technique in constructing scales measuring various aspects of attitudes to physics.

Both the sorting procedure and the rating procedure require the investigator to be clear in his own mind about what he is trying to measure. A test of such conceptual clarity is his ability to communicate his constructs to his panel of judges, and his ability to write items which reflect those constructs.

Whichever technique is used, factor analysis can then provide a further basis for allocating items to scales, for verifying the uniqueness of the various constructs, or for reducing the number of scales when some of them are shown to be redundant.[8]

Wilmut, like Selmes, also considers the question of whether attitudes to science, attitudes to scientists, and understanding of scientific method might be distinct variables. For him, the possibility is not so much dismissed as papered over, and a rather strange argument is put forward in the process:

> 'Clearly, these (attitudes) cannot be entirely separated and responses in one area will predetermine responses in another. For instance, a poor understanding of the methods of science might be expected to be reflected in a disagreement with the work of scientists and an emotionally unfavourable attitude towards science as the student experiences it in the school.' (1971, p. 83.)

Note, incidentally, the slide in argument from a strong and categorical assertion ('will predetermine') in the first sentence to a weak and probabilistic one ('might be expected to be reflected') in the second. Now, it may well be that poor understanding is associated with poor attitudes, but *this is a matter for empirical investigation, not untested assumption*. As a justification for adding together items to produce a single total score, it just will not do.

If empirical evidence were available to show that attitudes to science, attitudes to scientists and understanding are all strongly intercorrelated, then addition to form a single unidimensional scale would be justified. In the absence of such evidence, the variables should be measured separately and not added together.

Actually, two strands of evidence are available which suggest that these variables are *not* strongly intercorrelated, thus supporting the contention that they should be kept apart. The first strand is derived from the work of Jackson (1968) who has reviewed the literature on the relationship between the achievement of elementary school pupils and their attitudes to school 'Commonsense' might suggest that achievement ought to be correlated with attitude. Surprisingly, the correlations usually turn out to be exceptionally low. Thus, Wilmut's assumption that poor understanding might be expected to lead to emotionally unfavourable attitudes is highly speculative.[9]

The second strand of evidence comes from an investigation (Gardner, 1972) in which high school physics students were measured on four attitude variables: attitude towards discovery learning, attitudes towards the openness of physics (whether the students saw physics as flexible and dynamic, or static and unchanging), attitude towards scientists (eccentric or normal) and enjoyment of physics. Each of these attitudes was measured, needless to say, by a *separate* ten-item Likert-type scale; the items were randomly assorted within the 40-item instrument. The corrected split-half reliability coefficients for the four scales lay in the range ·75 to ·89; the six intercorrelations between the scales lay between ·24 and ·35. Thus, while the various attitudes were correlated with one another (e.g. pupils who enjoy physics *tend* to regard scientists as more normal), the fact that the intercorrelations are very much smaller than the reliabilities indicates that the variables are quite distinct. Factor analysis of the item responses provides further evidence of the uniqueness of the four scales. Here, then, is clear evidence that attitudes towards scientists and attitudes towards the study of a scientific subject are not the same thing.

The argument in this section has been complex and protracted. Yet all of it is nothing more than an elaboration of a simple theme: multi-dimensional entities should not be reduced to uni-dimensional variables, unless the reduction can be stoutly defended.

The Selmes scale is not the only one to attempt such a reduction. For example, in an instrument entitled *Attitude Toward Science* (Shah, 1962), some of the items reflect interest in science ('I like science as a hobby') while others reflect the respondent's attitudes to the significance of science in society ('I feel that science has grown up with the primary purpose of serving humanity'). Rational analysis ought to be sufficient to show that these are distinct variables which ought to be measured separately and not mixed together into a single score. Later research using other scales (Nuttall, 1971; Ormerod, 1971 and 1973) shows that these variables are correlated, but are certainly not identical.

A further example is provided by the work of Gaster (1972), who conducted an investigation of sixth-form science, mathematics and technology courses, and administered a number of instruments, including a *Science Motivation* scale. The scale consisted of three distinct sections. The first section presented a list of 26 occupations of which 14 were applied scientific/technical; respondents were asked to tick all that they felt would be interesting. The second section consisted of 12 multiple-choice cognitive items requiring students to demonstrate understanding of aspects of science in daily life (e.g., 'Which contributes most to the efficiency of a thermos flask? Glass/silvering/vacuum'. The third section ('hobbies and influences') contained eight assorted items asking about the respondent's interests, his father's occupation and other matters. Scoring procedures were not explicitly described, but apparently the single *Science Motivation* score was obtained by adding the number of scientific occupations ticked, the number of correct cognitive items and the number of science-based responses in the 'hobbies and influences' section.

The theoretical rationale for such a procedure is unclear. The researcher was interested in measuring creativity as a criterion variable; the *Science Motivation* scale, it is claimed, measures 'dominant motivational factors which might have influenced the subjects' performance in the creativity aspect of the battery'. Perhaps so, perhaps not; what is needed is an investigation of the properties of the scale. Simply combining cognitive and affective items and assuming that this combination measures motivation represents dubious psychometric technique.

A final illustration of Defect 2 is provided by the work of Bollen (1972), who devised an instrument to measure teachers' attitudes to the 'philosophy' of the Nuffield primary science course. The criticism here is much more gentle, since it is in many other respects a good scale: the items are well-written and there has been a clear attempt to relate their content to an explicit rationale.

The rationale consists of a number of value judgements held by developers of the Nuffield course. The key phrases in these value judgements are:

1. The teacher does not need to be a subject specialist....
2. He should not follow a set syllabus....
3. The children should be allowed to work in groups....
4. Investigation should be by practical means wherever possible....
5. The object . . . should be to promote an attitude of open-ended inquiry.
6. The teacher should have faith (in the approach)....

This collection forms a 'philosophy', but it does not reflect a single theoretical construct. There are at least two dimensions underlying this collection: belief in child-centredness, and belief in subject specialism. Statements 2 to 5 could probably all be regarded as aspects of the child-centred *vs.* the teacher-dominated dimension.

However, the first value judgement—reflected in the item 'A teacher must have a good scientific background'—hardly belongs on the same dimension. A teacher with a 'good Nuffield attitude' is supposed to disagree with the item. Yet it seems perfectly rational for a teacher to believe in a child-centred science curriculum and still believe that a good science background is helpful for teaching it. In other words, it is possible to believe that subject specialists make better teachers, or that child-centred education is valuable, or both, or neither. It follows that teachers' views about subject specialism and their views about child-centred education ought to be measured using separate scales; they are not reducible to a single score.

As it happens, the criticism just offered is perhaps somewhat pedantic. Most of the items on Bollen's scale reflect the child-centred/teacher-dominated dimension, and only one item reflects the subject specialist dimension. Most of the items therefore display internal consistency. However, the principle underlying the criticism is important: simply listing a set of value judgements and calling it a philosophy is not an adequate basis for defining a uni-dimensional construct.

Defect 3: Lack of relationship between treatment and outcome

A great deal of research in the social and medical sciences conforms to the model: apply treatment X and measure outcome Y. Library shelves have been filled with discussions of appropriate research designs and statistical procedures for dealing with experiments of this kind.

In any experimental situation, an unlimited number of treatments could be applied and an unlimited number of outcomes could be studied. How, then, does the researcher decide what to do? Frequently, but not always, he is guided by some sort of theory, perhaps sophisticated, perhaps primitive.[10] For example, if treatment X is 'add fluoride to town water supply' and outcome Y is 'incidence of tooth decay', the two apparently unrelated entities are connected by a theory which tells us how fluoride salts are incorporated into tooth enamel. If a researcher comes up with the outlandish proposal to study the effects of fluoridation upon the incidence of ingrown toe-nails, the onus is upon him (if he is not to be laughed out of court) to justify and explain why he thinks there might be a connection between the two.

Unfortunately, many educational experiments belong to the fluoridation/ingrown-toe-nails category. It is often less obvious that they do, because the investigator can frequently mask the lack of connection between treatment and outcome by the use of plausible language. For example, in the study by Wilmut (1971 and 1973), the treatment was 'Nuffield A-level physical science project work', and the outcome was 'attitudes to science'. This is not only plausible, but eminently reasonable—until one looks at the instrument used to measure the outcome.

At this point, it is appropriate to recall that a Likert-type scale is nothing more than the sum of its parts: unless the experimental treatment affects students' responses to a large number of the individual items, it will have no effect upon their attitude score. If the scale contains a large number of items which in no way relate to the experimental treatment, then it is highly unlikely that significant treatment effects will be found.

Again, this is all very obvious. Yet it is all too easily forgotten. The moment a respondent's total score has been calculated, it appears to assume a mystical significance of its own, transcending the mundane matter embodied in the items of which it is composed. Some of the items in the scale that Wilmut used have already been quoted. Here are five more:

'It is not possible to study human beings scientifically.'

'Women scientists are less feminine than other women.'

'The basic moral rule of a scientific society is simple; mutual respect, intellectual honesty, and good will.'

'The most important advances in science are made by a few outstanding men.'

'Scientists have proved that God does not exist.'

Since the research was specifically concerned with the effects of physical science project work, it is fair to examine each item in the criterion instrument carefully and ask, What is the connection between project work and responses to this item? Are the treatment and the measured outcome manifestly related? Theoretically related? Are the objectives of project work—which Wilmut discusses fully in an early chapter of his 1971 work—related to students' responses to these items? Does project work in physical science claim to alter students' views about the applicability of science to the study of human behaviour, about the femininity of women scientists, or about the fundamental morality underlying a scientific society? Is there a theory, even a primitive one, tying all these elements together?

One suspects that the answers to these questions are all a resounding no. Since the scale is composed of many more such items, it is hardly surprising that no significant differences were found between the experimental and the control groups. Partly, this is a result of the confounding of variables within the scale, criticized at length earlier. But not entirely: even if the criterion instrument were divided into separate, well-defined scales, the question of the relationship between treatment and outcome would still remain.

In the light of this discussion, it is perhaps pertinent to note that scale validity is not a property inherent in a scale: it depends very much upon the purpose for which the scale is used. A spatial relations test may be valid for measuring the effects of play with blocks and construction kits, but quite invalid for measuring the effects of an audio-visual language course. Dr A may construct an attitude scale which he claims is valid, but this does not mean that Mr B should use it in studying treatment effects unless the theoretical construct being measured by the scale can be shown to be related in some way to the treatment. Again, all very obvious. Yet some educational researchers are still using scales apparently because they are *available* without first deciding whether they are *appropriate*. A much better procedure would be:

(1) Read the writings of curriculum developers, analyse instructional materials, interview parents, pupils, headmasters and teachers, in order to discover implicit or explicit objectives held to be valuable.

(2) Translate these objectives into clearly defined constructs, and devise scales (if they are not already available) to measure these constructs.

(3) In studies of the effects of experimental treatments, restrict the outcomes measured to those which can be related in a logical way to aspects of the treatment.

Conclusion

Three fundamental principles have been stressed throughout this paper. Measurement should reflect clear concepts. Measurements should not be combined without reason. In experiments, the measured outcomes should display some defensible connection with the applied treatments. They are all rather obvious. Then why are they so frequently ignored? It is almost as if the inner urge to get to work with a calculating machine or computer is so strong—look at me, I can do analysis of variance!—that the question of what the numbers actually *mean* is set aside. It is perhaps time that some emphasis was placed on first principles. It is perhaps time for university departments which encourage their post-graduate candidates to engage in number-stuffing in the misguided belief that they are being trained to do educational research to be rapidly and bluntly disabused of the notion.

Notes

1. Of course, it is not always a simple matter to define and measure psychological concepts. Consider, for example, the concept of *curiosity*. There is a growing body of research which is attempting to measure curiosity in young children: Berlyne (1966) and Hutt (1970) provide good examples. The work of both of these researchers is characterized by a serious attempt to define and operationalize the concept of curiosity.

2. Some of the studies have appeared as published work; others are unpublished theses. Most educational research is probably done for higher degrees, so university theses are by no means unrepresentative of research generally. In addition, it is often only in detailed theses that there is sufficient information to judge the quality of the measuring instruments employed.

3. Harrison reports that the split-half reliability coefficient of the instrument was ·63. This is extremely low, considering that there were 50 items in the instrument. (The reliability of a scale is most strongly affected by the number of items in it.) The low value faithfully reflects, of course, the lack of internal consistency in the items.

4. A third instrument, the *Science Opinion Poll* (Laughton and Wilkinson, 1968) provides additional illustrative material for this paper; however, their report does not include a listing of the items employed, and it is therefore difficult to decide whether their scale illustrates Defect 1 or is better classified as an example of Defect 2. Some idea of the item content may be gained from the following description of the instrument:

   '. . . the Science Opinion Poll has been devised by the authors to examine the attitudes to the study of science of pupils following different courses. The test is not measuring the extent to which the objectives of science courses have been achieved but only whether or not pupils regard these objectives as important. The test, we hope, is also a measure of a pupil's interest in science.

   Seventy items formed the initial draft of the test. These consisted of statements expressing a favourable or an unfavourable opinion about the study of science. . . . These statements covered pupils' attitudes to various aspects of the study of science of which the following are examples—practical work; science as a subject in the curriculum; characteristics of scientific thinking such as honesty, planning experiments, interpretation of evidence, etc; interest and enjoyment in science . . .'

   At best, this is a confused mixture of a number of separate scales, and would illustrate Defect 2; however, one might also argue that there is no well-defined theoretical construct underlying the item content, and this would be an example of Defect 1. In passing, it might be noted that the study displays other weaknesses, notably a total absence of control over extraneous variables inextricably confounded with the treatment variable.

5. It is debatable whether this third category is predominantly an attitudinal trait.

6. One might wonder what happened to the attributes in Diederich's list. *Are* attributes like scepticism, faith, aversion to superstition and all the rest essential components of scientific attitudes? Are they perhaps incorporated within the broader headings of Aiken and Aiken's list? If not, why were they omitted?

7. To take an exact parallel from the cognitive domain: consider an 'intelligence' test containing 50 spatial visualization items and 50 vocabulary items. One item differentiates amongst 8-year-olds (brighter 8-year-olds get it right, duller 8-year-olds get it wrong) but does not differentiate amongst 5-year-olds (most get it wrong). This information is quite unrelated to the question of whether the test should be scored as a 100-item 'intelligence' test, or as two separate 50-item tests measuring spatial and verbal skills.

8. However, it should be noted that factor analysis is a refining and not a transforming procedure: it can turn brown sugar into white sugar, but it cannot make silk purses out of sows' ears. For the educational measurement worker, all the statistics in the world are no substitute for good hard preliminary thinking about what it is that he is trying to measure.

9. Wilmut is not the only attitude researcher to make this assumption. The *Science Interest* scale of the NFER *Science Attitude Questionnaire* (Skurnik and Jeffs, 1970) contains the following item:

   'I do very badly in science'

   This item indicates the student's perception of his own achievement, but it is hardly an indicator of his interest in the subject. To include such an item is to make the same unexplored assumption about the relationships between attitudes and cognitive variables that Wilmut is making.

10. Theories, even primitive ones, are not always used as guides to research. For example, in many drug trials, something akin to trial-and-error is frequently employed. The pay-offs (lives saved, profits for drug companies) make such a strategy worthwhile. The pay-offs of educational experiments are seldom so dramatic as to warrant trial-and-error strategies.

References

Aiken, L. R. Jr. and Aiken, D. R. (1969). 'Recent research on attitudes concerning science', *Science Education*, **53**, 295–305.

Berlyne, D. E. (1966). 'Curiosity and exploration', *Science*, **153**, 25–33.

Bollen, F. A. (1972). 'Attitude assessment in science teaching', *School Science Review*, **53**, 217–35.

Diederich, P. B. (1967). 'Components of the scientific attitude', *Science Teacher*, **34**, 23–4.

Gardner, P. L. (1972). 'Attitudes to physics'. Unpublished PhD thesis, Monash University.

Gaster, E. J. 'The measurement of some aspects of integrated science, mathematics and technology courses at sixth form level.' Unpublished MPhil thesis, University of Nottingham.

Genn, J. M. (1970). 'Chemistry scholars in school and university contexts: a study of transition from school to university.' In: Campbell, W. J. (Ed.) *Scholars in Context*, Sydney: Wiley.

Harrison, I. V. (1971). 'Development of an attitude test for Ordinary National Certificate students in sciences'. Unpublished MSc thesis, Bradford University.

Hutt, C. (1970). 'Curiosity in young children', *Science Journal*, **6**, 2, 68–71.

Jackson, P. W. (1968). *Life in Classrooms*. New York: Holt, Rinehart & Winston.

Laughton, W. H. and Wilkinson, W. J. (1968). 'Pupils' attitudes to science teaching', *Education in Science*, February, 31–3.

Nuttall, D. (1971). *Administrator's Manual for Science Attitude Questionnaire*. Slough: NFER.

Ormerod, M. (1971). 'The "social implications" factor in attitudes to science', *Brit. J. Educ. Psych.* **41**, 335–8.

Ormerod, M. B. (1973). 'Social and subject factors in attitudes to science', *School Science Review*, **54**, 654–60.

Selmes, C. S. (1971). 'Attitudes to science: the design, construction, validation and use of an attitude scale.' Unpublished PhD thesis, University of Bath.

Shah, S. A. (1962). 'A study of the attitudes of students towards the science course at certain training colleges in the United Kingdom and at a training college in Trinidad'. Unpublished MA thesis, University of London.

Skurnik, L. S. and Jeffs, P. M. (1970). *Science Attitude Questionnaire*, Slough: NFER.

Wilmut, J. (1971). 'The effect of project work on pupils' attitudes to science'. Unpublished MEd thesis, University of Bath.

Wilmut, J. (1973). 'The effect of project work on the attitudes to science of sixth-form pupils'. *Educ. Res.* **15**, 2, 128–33.

Journal of Curriculum Studies **Collins**

In the May issue

Curriculum Criticism and Literary Criticism.
George H. Willis

Music Education: Some Critical Comments.
Graham Vulliamy

A Learning Hierarchies Approach to Social Studies.
Basil S. Deming

Group Work with Early School Leavers. *David B. Clark*

Secondary Music Teachers' Attitudes towards the Teaching of Their Subject. *Lyndon Jones*

A Study of the Stability of Biology-teachers' Priorities and Expectations. *E. Jungwirth*

Readability Indices in the Evaluation of Curriculum Materials. *E. Ayotunde Yoloye*

It may be ordered direct from the Publishers

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada
APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

May and November
Annual Subscription £3.00 including postage

Attitudes of Nigerian Secondary School Children to Programmed Instruction in Geography

P. O. Okunrotifa, Department of Education, University of Ibadan

Summary

This paper reports an investigation into the attitudes of 200 mixed secondary pupils from both rural and urban schools towards programmed instruction as a teaching method and towards geography, as a subject studied in this way. Pupils in their second year in each school were divided into experimental and control groups matched on the basis of achievement, verbal and quantitative aptitude tests. The attitudes of both groups towards geography and programmed instruction were pre-tested by questionnaire. The groups then underwent two different learning sessions for three weeks. Despite more positive attitudes shown initially by the control group to both geography and programmed learning, the experimental group later performed better and showed more favourable attitudes towards map reading than those taught by conventional texts, this being even more so for boys. Reasons for the results and certain features of the study are discussed.

Introduction

Psychological literature is replete with data attesting to the general effectiveness of learning from programmed instruction. Even in Nigeria where only a very limited number of studies in the field has been undertaken, Roebuck (1967) and Okunrotifa (1968) have used British and American programs in Nigerian conditions to indicate that relevant self-instructional programes could promote learning. However efficient programmed instruction may appear to be, the fact still remains that educators cannot assume that the attitudes of their pupils will be favourable towards the new instructional device without some research results to support such a claim. How valid is the assumption that Nigerian children will react favourably towards the kind of thinking and learning which is called for in using programmed instruction? What will be the attitudes of these children towards geography when it is presented through programmed instruction? It is the purpose of this paper to find answers to such questions. Specifically, the hypotheses tested were as follows:

Ho_1 There would be no significant difference in the attitudes to geography between

(a) the two treatment (programmed and conventional text) groups;
(b) the two sexes—male and female;
(c) the two environment (urban and rural) groups;

in the study.

Ho_2 There would be no significant difference in the attitudes to programmed instruction between the two

(a) treatment groups;
(b) sex groups; and
(c) environment groups;

in the study.

Subjects

On the bases of area, population size, functions and position of the settlements in the North Central State of Nigeria where the present study was carried out, four schools representing urban boys', urban girls', rural boys' and rural girls' schools in the state were randomly selected for the study. A total sample of 200 second formers (eighth graders), 100 boys and 100 girls was used. The mean age of the subjects was fourteen years. None of the 200 subjects had had previous exposure to programmed instruction materials.

Procedure

Two types of research instruments were used in this study: learning materials and measuring instruments. The learning materials consisted of five geography programmed texts, five conventional texts in geography covering the same span of information as the programs, and a familiarization program in civics. The measuring instruments also consisted of two aptitude tests—verbal and quantitative—obtained from the West African Examinations Council, a geography achievement test, and also two Likert-type attitude questionnaires to measure attitudes towards map reading and programmed instruction. The development and validation of all the research instruments have been described elsewhere (Okunrotifa, 1973).

During the first three experimental sessions, all subjects were taught about certain methods of learning geography. They were taught with the aid of a familiarization program in civics about programmed instruction and how to use linear programs. In each school, subjects were divided into experimental and control groups. The two groups were matched on the bases of pre-test geography achievement, verbal aptitude and quantitative aptitude tests. Both groups were given pre-test attitude towards geography and pre-test attitude towards programmed instruction (Likert-type). These were followed in each school by fifteen experimental sessions (days) extending over three weeks during which the experimental subjects learned from programs while the control group learned from conventional texts. The topic: map reading in geography, served as the common treatment component. Post-treatment measures of the relevant variables were again obtained and sub-groups were compared on pre-test and post-test changes.

In order to control possible methodological errors during the study, subjects were informed at the first session that the experiment was concerned with student learning in geography. To enlist their co-operation, it was made clear that leakage could destroy the experimental test of the independent variables. However, subjects were not told that different methods of teaching would be employed and that a certain class would be designated as experimental group or as control group. Subjects were not also made aware of the investigator's hypotheses.

Experimental bias known as the 'Rosenthal Effect' (Rosenthal, 1964) was adequately controlled since all sessions were conducted by two other experimenters who were not aware of the particular hypotheses being tested. During experimental sessions, the investigator maintained an unobtrusive presence as far as the subjects were concerned and came in only to answer questions about procedural matters. Conditions of uniformity were established as far as possible for working on the programmed and conventional texts. Each was carried out on the same week-day, during the same lesson period. The 'Hawthorne Effect', it was believed, would cancel out since both groups were being tested and the investigator and the experimenters were not normally teachers in the school. However, not so much success was achieved with the control of certain factors outside the experiment or in the general school situation. Thus, such factors as concern for the impending August school examination, and subjects' personality differences were not effectively controlled and they might have contributed to reduce the effectiveness of the experimental manipulation.

Results and discussion

The assumption of initial equality of groups was first tested and the results are presented in Table 1. There were no significant differences in verbal aptitude, quantitative aptitude, pre-test geography achievement (the matching variables) and age. Significant differences, however, existed in pre-test geography attitude and pre-test program attitude—the control group having more positive attitudes in both cases. The significant differences indicated the need for statistically equating experimental and control groups on the two variables before making further comparisons concerning attitudes towards geography and programmed instruction.

A description of the performances of the control and experimental groups in geography has been reported elsewhere (Okunrotifa, 1974), but a summary of it is given here. Treatment groups were compared for geography achievement by a three-way analysis of covariance with teaching methods, sex and school environment as main effects using pre-test scores as covariate

Table 1: Similarity of experimental and control groups, pre-test

	EXPERIMENTAL GROUP (N=100)		CONTROL GROUP (N=100)		
Variable	*Mean*	*SD*	*Mean*	*SD*	*t*
Age	14·21	1·05	14·26	·31	−·36 (NS)
Verbal aptitude[a]	37·56	7·62	37·22	7·37	·32 (NS)
Quantitative aptitude[b]	30·57	9·46	30·36	8·82	·16 (NS)
Pre-test geography achievement[c]	35·63	9·45	35·90	9·92	−·20 (NS)
Pre-test geography attitude	83·78	10·20	86·81	12·59	−·2·06*
Pre-test program attitude	136·24	21·44	141·99	19·51	−·2·22*

a, b, c, Matching variables *p< ·05

Table 2: Experimental and control group means on post-test geography attitude scores adjusted for pre-test scores

	EXPERIMENTAL GROUP				CONTROL GROUP			
School	*n*	*X-Means*	*Y-Means*	*Adjusted Y-Means*	*n*	*X-Means*	*Y-Means*	*Adjusted Y-Means*
URBAN	*60*	*85·58*	*98·20*	*98·03*	*60*	*86·81*	*92·39*	*91·50*
Male	30	86·87	98·70	97·77	30	92·73	98·40	94·02
Female	30	84·30	97·70	98·29	30	84·80	90·70	91·00
RURAL	*40*	*81·08*	*92·93*	*95·42*	*40*	*83·88*	*89·15*	*89·99*
Male	20	82·35	95·70	97·44	20	86·05	90·95	90·51
Female	20	79·80	90·15	93·40	20	81·70	87·35	89·47

Table 3: Analysis of covariance for post-test geography attitude scores adjusted for pre-test scores

Source of Variation	*DF*	*Sum of Squares*	*Mean Square*	*F*
Teaching Methods	1	1469·48	1469·48	16·19**
Sex	1	153·22	153·22	1·69 (NS)
School Environment	1	307·25	307·25	3·38 (NS)
Interaction	1	159·53	159·53	1·76 (NS)
Within	194	17612·24	90·78	

** p<·01

and post-test scores as criterion. There was a significant difference in teaching methods. An F-ratio of 23·56 revealed a significant treatment difference (·01 level) in favour of the programmed group and there were no significant differences found among the other variances. A 2×2×2 factorial analysis of variance was also computed for the comparison of groups' learning times scores. There were no significant main effects for methods, sex and school environment nor were there any significant interactions involving sex, methods, and school environment variables. On the basis of these results, it was therefore concluded that programmed instruction was more efficient than conventional instruction in contributing to pupils' achievement.

In testing the first hypothesis of the present study, groups were compared by using a three-way analysis of covariance with methods, sex and school environment as main effects. The pre-test attitude towards geography was used as a control variable and the post-test attitude towards geography served as the criterion variable. The results of the analysis of covariance contained in Tables 2 and 3 show that method variances were significant while the other variances were

Table 4: Experimental and control group means on post-test program attitude scores adjusted for pre-test scores

	EXPERIMENTAL GROUP				CONTROL GROUP			
	n	*X-Means*	*Y-Means*	*Adjusted Y-Means*	*n*	*X-Means*	*Y-Means*	*Adjusted Y-Means*
URBAN	*60*	*138·77*	*149·50*	*149·66*	*60*	*145·33*	*137·77*	*134·85*
Male	30	145·60	155·03	151·98	30	152·33	144·77	138·56
Female	30	131·93	143·97	147·35	30	138·30	130·77	131·16
RURAL	*40*	*132·45*	*143·05*	*146·18*	*40*	*136·98*	*129·88*	*130·89*
Male	20	139·05	149·15	149·18	20	138·90	130·65	130·75
Female	20	125·85	136·95	143·19	20	135·05	129·10	131·01

Table 5: Analysis of covariance for post-test program attitude scores adjusted for pre-test scores

Source of Variation	*DF*	*Sum of Squares*	*Mean Square*	*F*
Teaching Methods	1	11036·11	11036·11	43·12**
Sex	1	1023·57	1023·57	3·99*
School Environment	1	637·20	637·20	2·49 (NS)
Interaction	1	368·18	368·18	1·44 (NS)
Within	194	49653·69	255·95	

* $p<\cdot05$ ** $p<\cdot01$

not. This means that the programmed instruction group, irrespective of differences due to sex or school environment, showed a more significantly favourable attitude towards map reading than the conventional text group.

The more favourable attitude of the programmed group could be attributed to two factors. First, programs, like the conventional texts, were well validated and so were successful as teaching instruments. No doubt, a poor program, resulting in unsuccessful learning, could probably have led to poor attitudes. Second, the fact that some major principles of learning—such as immediate knowledge of results, active response, constant evaluation, appropriate practice and graduated sequence—are usually more emphasized in programmed instruction than in conventional text instruction might have made map reading easier and more satisfying and therefore more liked by the programmed group. The positive and significant correlation ($^{r}57\cdot234 = 0\cdot198$)[1] between post-test geography achievement and post-test geography attitude appears to reinforce this view. However, these explanations are purely speculative. It is recognized that the reasons for the overwhelmingly favourable response from the experimental subjects may be difficult to evaluate. However, since such transient factors as novelty effect or 'Hawthorne Effect' were likely to have affected both treatment groups equally, it would be concluded that a high degree of the intrinsic value in programmed method contributed to the result.

A 2×2×2 analysis of covariance was also used to test the second hypothesis of the present study. Tables 4 and 5 show that when post-test program attitude scores were adjusted for pre-test program attitude scores, the F-ratios turned out to be significant only for variances associated with methods and sex variables. This means that the group which actually learned from programs developed a more significantly favourable attitude towards the new instructional device than the conventional text group which did not learn from programs. Furthermore, there were significant differences between sex groups because boys displayed more favourable attitudes towards programmed instruction.

A significant point worthy of note is that whereas the control group had more positive attitudes initially (see Table 1) than the experimental group, Table 4 clearly shows that while the attitude mean score changes for the experimental group were in positive direction, those of the control group were in negative direction. Thus, it would seem that while the experimental subjects developed a more favourable attitude towards programs, the control subjects' attitudes declined from an initial enthusiasm to moderately positive reactions.

---

[1] With verbal (2) and quantitative (3) aptitude tests and pre-test achievement (4) partialled out, the original correlation, $^{r}57 = 0\cdot282$ (in the experimental group) between post-test geography achievement (5) and post-test geography attitude (7) decreases to a partial coefficient, $^{r}57\cdot234 = 0\cdot198$ which is significant at the ·05 level.

An important factor which might have influenced attitude differences between treatment groups was the varying length of contact which experimental and control groups had with programs. The control subjects learned about programs during the three introductory sessions only. During the learning sessions, they learned from conventional texts. On the other hand, after the introductory sessions, experimental subjects really learned from programs throughout the learning sessions and so they had more opportunity than the control subjects for a critical assessment of the nature of programs. The learning experiences derived from such a critical assessment might have proved satisfying and so contributed to a more favourable attitude than in the control group where the length of subjects' contact with programs was much shorter and limited. However, the low correlation ($r = \cdot 142$) found between post-test attitude towards programs and post-test geography achievement in the experimental group may suggest that a positive attitude towards programmed instruction does not of itself necessarily result in high achievement. It would seem that, in contrast to attitudes towards geography, pupils' attitude towards programs was not an important variable in the programmed learning situation of the present study.

Another result obtained from a test of the second hypothesis is the significant sex difference in attitudes towards programmed instruction. Boys liked programs more than girls. The more favourable attitude of the boys could possibly be attributed to the fact that programmed instruction calls for some mechanical reasoning and according to Anastasi (1967), boys are superior in mechanical aptitude to girls.

Conclusion

The present study indicates that, irrespective of differences in the sex or school environment of learners, validated programs can produce more favourable attitudes towards map reading in geography than well validated conventional texts covering the same span of information as the programs. This implies that programmed instruction may be considered as one of an arsenal of teaching methods at the command of the teacher to help him to promote efficient learning and a favourable attitude towards the school curriculum.

References

ANASTASI, A. (1967). 'Mechanical aptitude', in *Psychological Testing*. New York: Macmillan, p. 387.

OKUNROTIFA, P. O. (1968). 'A comparison of the responses of Nigerian pupils to two sets of programmed instruction materials in Geography', *Programmed Learning and Educational Technology*, Oct., 283–93.

OKUNROTIFA, P. O. (1973). 'The establishment of the validities of experimental materials in Geography', *The Journal of Geography*, Nov. Washington: National Council for Geographic Education.

OKUNROTIFA, P. O. (1974). 'The performance of Nigerian pupils in a programmed unit of Geography', *Savanna*. Ahmadu Bello University: Zaria, Nigeria.

ROEBUCK, M. (1967). 'A report on the evaluative use of the program "Bracket" in Form 2 of the International School, Ibadan'. Institute of Education, University of Ibadan.

ROSENTHAL, R. (1964). 'The effect of the experimenter on the results of psychological research', *Progress in Experimental Personality Research*, *1*, 79–114.

The Role of the Physical Education Teacher

L. B. Hendry, Department of Education, University of Aberdeen

Summary

A survey of the expectations of physical education teachers, classroom colleagues, head teachers and pupils in 13 non-selective urban schools serving one college of education, and of students attending a specialist physical education college revealed both concurring and conflicting view-points. The findings suggested that physical education teachers have low subject status and élitist values, and that classroom colleagues provide role reinforcement for the teacher of physical education in a situation where there are differing expectations about his role. It was also shown that students in training quickly acquire an understanding of their future teaching role.

Introduction

The present paper attempts to consider the physical education teacher's role in terms of (a) students' expectations across the years of teacher training; (b) how pupils see the role of the physical education teacher; and (c) the perceptions (and possible reinforcing influences) of classroom colleagues.

Teacher training and pupil-relations (a and b)

Cohen (1967) considered that courses of professional training are designed on the assumption that students gradually immerse themselves in the teacher's role by means of college studies, by experiences with children, and by encounters with teachers. Physical education students may well be selected 'in the image'—attitudinal as well as physical—of those carrying out selection for teacher training (Carter, 1964; Ward, 1968). There is the possibility of a closed system where certain traditions and values are retained and perpetuated. This could be demonstrated by showing that when these highly skilled pupils become teachers themselves they continue to reinforce a success syndrome rather than being seen to cater for the needs and interests of a wide range of pupils. Obviously the physical education teacher, involved in a non-examinable subject, does not necessarily have to ensure a specific form of achievement, but can encourage success in many different areas of physical activity. In a recent article, however, Reid (1972) has posed a provocative question in evaluating pupils' involvement in extra-curricular activities: 'Why don't a higher percentage of the able pupils belong to activities other than the "prestige" activities?' These 'prestige' activities were either sports teams or school choirs, and it was suggested that in some schools at least the integrating function of such activities for pupils may be sacrificed to achieve a higher level of competitive performance for the chosen few. Biglin (1968) has also pointed out that it seems reasonably safe to assume that where there is diversity between the physical education teacher's role conceptions and the pupils' viewpoint, some kind of conflict may result. Thus situations can occur where many pupils are not particularly attracted to physical education lessons and teachers concentrate on the gifted few.

Role expectations and school functions (c)

Physical education students are aware of the low status of their subject in schools (Appleby, 1967). Yet a higher percentage of students attending specialist physical education courses have university entrance qualifications than students following other college courses (Hargreaves, 1968). The suggestion here is that it is the nature of the subject and its perceived values within the school system which influence status rather than the qualities of the occupant of the particular teaching role.

Physical education students and teachers have indicated that their main educational priority lies in contributing to the social development of pupils (Lee, 1968; Heaton, 1969), and so clearly there is no direct involvement for the physical education teacher in the academic or vocational main-stream of the schools' functioning. In addition, the physical education teacher's educational medium is teaching physical skills. Hence, Musgrove and Taylor (1969) have argued that the physical education teacher is seen as carrying out a 'trainer-instructor' function rather than viewed as a 'teacher'. The public image of the physical education teacher has been described by Scotland (1964) as 'generally a man of action rather than mind. A good man to play golf with, but not first choice for a discussion of professional problems.'

The present study is concerned with expectations for the physical education teacher's role from the viewpoint of students during their teacher training course, of classroom colleagues, head teachers and pupils, as well as the expectations of practising teachers of physical education. In particular it attempts to examine the development of role in students, the role orientation of physical education teachers themselves, and the possible conflicting and reinforcing influences in the school setting.

Procedure

A. Sampling

Thirteen out of 16 non-selective schools in the urban area of one college of education in the North of England agreed to co-operate in the survey.

In each participating school the following groups completed a questionnaire on the role of the physical education teacher:

(a) male physical education teachers with main course or specialist qualifications (N=25) and, in addition, 9 teachers from the nine selective schools in the college's teaching practice area (N=34);
(b) Five male classroom teachers randomly selected from the staff list (though some 22 per cent declined to complete the questionnaire) (N=51);
(c) headteachers (N=13);
(d) Five male pupils[1] aged 14 years randomly selected from the school year list (N=68).

In addition, 39 out of 40 third-year and all first-year male students from a specialist physical education college in the same geographical area also completed the questionnaire:

(a) first-year physical education students (N=41);
(b) third-year physical education students (N=39).

B. Construction of the questionnaire

The role questionnaire was constructed with the help of experienced physical education teachers, lecturers, psychologists and sociologists. Initially 140 statements were devised to cover seven areas of the physical education teacher's role, characteristics and motivations, namely:

(a) status within the school;
(b) relations with other staff;
(c) relations with the head teacher;
(d) relations with pupils;
(e) extra-curricular duties;
(f) personal characteristics;
(g) conscious motives for career selection.

These 140 statements (i.e. 20 items × seven role areas) had been presented to a group of 25 qualified and experienced physical education teachers[2], who were asked to select (independently) the five most important statements in rank order for each of the seven areas. Each item chosen as first in importance by respondents was given a weight of 5, the next 4, and so on to 1, and the final version of the adults' questionnaire was constructed by selecting the 40 items (five or six items from each of the seven role areas) with the largest weighted scores. The presentation of statements in each area of the questionnaire was randomized to avoid patterning of response.

Pupils' expectations were assessed in only two of the seven selected areas of the adults' questionnaire (namely 'relations with pupils' and 'personal characteristics'). Because of this it was decided to examine significant shifts of opinion (Brown, 1956) from 'actual' to 'ideal' role conceptions based on responses to the 36[3] items contained in the pupils' questionnaire. These role terms were used in the meaning adopted by Westwood (1967):

---

[1] By mistake, one school had eight pupils complete the questionnaire.

[2] These teachers took no further part in the investigation. The deliberate choice of PE teachers in selecting items for inclusion in the final version of the questionnaire was made in an attempt to examine differing definitions of the PE teacher's role.

[3] Only a few items were removed from the original statements (20 items x two) comprising the pupils' questionnaire.

(a) how the individual probably will behave—'actual' role conceptions;

(b) how the individual should behave—'ideal' role conceptions.

Questionnaires were completed by pupils under group examination conditions and supervised by a member of staff not involved in the survey. Students were supervised under similar conditions. Teachers completed their questionnaires privately. In all cases complete anonymity was preserved.

Reliability coefficients for the questionnaire were obtained with a group of 20 third-year students and yielded a test-retest result of +0·82 over three months.

Results

Concurring view-points between groups on various role areas are presented descriptively in the text. Only significantly different expectations from questionnaires are presented in tabular form.

Results from the role questionnaires showed areas of both consensus and disagreement in role expectations.

Table 1: Significant differences between physical education teachers' role conceptions and students' expectations

Item No.	Questionnaire Items on the Male PE Teacher	PE Teachers N=34*		3rd-Year Students N=39		1st-Year Students N=41		†	
		A	D	A	D	A	D	χ^2	ρ
2	Is less important as time-table becomes 'exam-oriented'	20	12	13	25	16	24	6·2	0·05
4	Is judged on the school's sporting successes	26	8	30	7	23	18	6·7	0·05
5	Work not expected to interfere with academic work	25	8	21	17	12	29	16·1	0·01
6	Lessons most likely to be interrupted for interviews (form filling by pupils, etc.)	14	20	26	10	25	16	7·1	0·05
7	Is regarded as lacking in intellect and as non-academic by other staff	13	21	28	10	18	21	10·2	0·01
9	Does not 'sell' his subject well, so his role is misunderstood	20	13	27	13	12	22	8·3	0·05
14	Has smaller responsibility allowances in salary	29	4	16	12	18	8	7·4	0·05
18	Has *especially* good relations with pupils	29	5	36	2	25	16	14·7	0·01

A=Agree; D=Disagree: df 2.

* No significant differences in pattern of response were found between the nine teachers from selective schools and the 25 teachers from non-selective schools.

† Separate analyses by Chi-square (df 1) (Teachers *v.* 3rd-year students; Teachers *v.* 1st-year students; 1st-year students *v.* 3rd-year students) showed a trend towards alignment of PE teachers' responses with 3rd-year students in items 4, 5, 9, 18; whereas the other four items indicated differences in response between practising teachers and one or other of the groups of students.

1. *Students' Expectations*

(a) Physical education students and practising teachers were fairly similar in their expectations. Their responses showed that they tended to agree about the social function of the subjects, that it was easy to have good relations with pupils, and about the low status of the subject within the school. (There was agreement in 32 out of 40 items).

(b) The differences which did emerge tend to support the explanation which Cohen (1967) has given of how role conceptions are acquired. This process whereby students across their training years align their views to those held by teachers he has termed 'role accretion' (see Table 1, items 4, 5, 9, 18). Other significantly different items suggested varying perspectives of teachers and students about the social factors affecting the status of physical education in the school.

2. *Pupils' expectations*

All pupils considered that the physical education teacher had a strong influence on their personality development, and the majority of pupils thought that he was the kind of teacher a boy could confide in, a pupil-oriented social leader. Eighty per cent agreed with his disciplinary function in school, and the same number said he should be involved in counselling pupils.

There were significant shifts of opinion between the pupils' statements of how they expected the physical education teacher to behave ('actual role') and how they thought he should behave ('ideal role'). The physical education teacher was perceived by some as giving more attention to the more highly skilled pupils. Thirty-eight per cent indicated that the physical education teacher tended to concentrate only on the very able and talented pupils, and showed some lack of sympathy and understanding of the less able and interested. Responses to 'ideal' role expectations showed that approximately 96 per cent of pupils held the view that all pupils should be given attention, sympathy and understanding; there was also a significant shift of opinion to indicate that ideally a majority of pupils felt that the physical education teacher should have some influence on pupils' future careers. These three shifts of opinion from 'actual role' to 'ideal role' were significant at the 5 per cent level.

3. *Colleagues' and head teachers' expectations*

(a) Colleagues tended to hold similar views to physical education teachers about why these specialist teachers got smaller responsibility allowances. Reasons such as the head teacher's perceptions of low subject status, and of few opportunities given for general responsibility within the school, were presented. The physical education teacher's prestige was seen to be gained by the success and achievement of sporting teams. Colleagues accepted that the physical education teacher was able to offer different insights on pupils' character (97 per cent) but they indicated that he was not fully able to present coherently the values of the subject. They acknowledged that the physical education teacher appeared to have good relations with pupils, and that he exerted an influence on pupils' moral and social development. Colleagues considered that part of his duties included extra-curricular teaching; and that he needed to maintain friendly staffroom relations to ensure help with these extra-curricular activities. Eighty per cent of colleagues indicated that the physical education teacher was particularly friendly with young members of staff and 'those interested in sport'.

(b) There was agreement between head teachers' expectations and those of physical teachers about the physical education teacher's disciplinary influence in the school; that he could provide information about pupils from a different viewpoint; and that he had particularly good relations with, and influence on, pupils.

The physical education teacher was also seen as a devoted member of staff with a number of highly acceptable personal and professional qualities. Seventy-seven per cent of head teachers and 85 per cent of physical education teachers agreed that the physical education teacher maintained good relations with staffroom colleagues to ensure extra-curricular help, and all head teachers considered that extra-curricular involvement was an aid to promotion for the physical education teacher.

(c) Physical education teachers' expectations differed significantly from those of both colleagues and head teachers in certain areas of the physical education teacher's role performance. These differences centred on claims by physical education teachers that they showed greater efficiency in carrying out their teaching role, spent more time on the job or showed greater dedication than other teachers. Ninety per cent of colleagues and all head teachers did not agree with these claims.

(d) Seventy-six per cent of physical education teachers considered their work should not interfere

Table 2: Significant differences between physical education teachers' role conceptions and head teachers' expectations

Item No.	Questionnaire Items on the Male PE Teacher	PE Teachers N=34 A	PE Teachers N=34 D	Head Teachers N=13 A	Head Teachers N=13 D	χ^2	p
2	Is less important as time-table becomes 'exam-oriented'	20	12	1	12	9·1	0·01
3	Is expected to work with more pupils per class than other staff	26	6	4	9	8·5	0·01
4	Is judged on school's sporting successes	26	8	0	13	*	0·01
5	Work not expected to interfere with academic work	25	8	0	11	*	0·01
7	Is regarded as non-academic and lacking in intellect by other staff	13	21	0	13	*	0·01
13	Is considered by head teacher as taking second place to academic subjects, and has low status	25	9	1	10	11·6	0·01
15	Gets on well with headmaster only if he presents no difficult professional demands	18	5	1	12	13·9	0·01
24	Spends more free time on his job than otber staff	29	5	6	7	5·7	0·05
25	Is the most devoted member of staff	14	16	0	12	*	0·01
26	Spends more free time with pupils than other staff	30	4	7	6	4·8	0·05

A=Agree, D=Disagree.
* Denotes significance by Brown's (1956) formula.

with academic studies. But no head teacher supported this view, and only half the classroom teachers agreed with it. All headteachers disagreed that the subject's prestige hinged on the success of sporting teams. While physical education teachers reported subject marginality in relation to classroom subjects (85 per cent), and the need for conformity in their relations with the headteacher (80 per cent), headteachers strongly disagreed.

(e) Colleagues revealed uncertainty about the physical education teacher's disciplinary function but 85 per cent of head teachers and 83 per cent of physical education teachers considered this as part of the physical education teacher's role.

Discussion

Basically, the findings of the present study have highlighted four main points.

Firstly, physical education students appeared to have an early general understanding of their future

Table 3: Significant differences between physical education teachers' role conceptions and classroom colleagues' expectations

Item No.	Questionnaire Items on the Male PE Teacher	PE Teachers N=34 A	PE Teachers N=34 D	Colleagues N=51 A	Colleagues N=51 D	χ^2	ρ
1	Is used as an influence of discipline in school	28	6	26	25	7·4	0·01
3	Is expected to work with more pupils per class than other staff	26	6	14	37	20·7	0·01
5	Work not expected to interfere with academic work	25	8	25	25	4·5	0·05
15	Gets on well with headmaster only if he presents no difficult professional demands	18	5	19	26	6·6	0·05
24	Spends more free time on his job than other staff	29	5	30	20	5·0	0·05
25	Is the most devoted member of staff	14	16	4	46	13·9	0·01
26	Spends more free time with pupils than other staff	30	4	34	17	4·0	0·05

A=Agree, D=Disagree.

role, namely, a social orientation in their teaching and relatively low status, as has been described by other researchers (Appleby, 1967; Lee, 1968). Their early understanding of role may be partly explained by the fact that they were the successful pupils within the school physical education programme. 'Role accretion' apparently occurs for students of physical education in much the same way as for other students, and this may be considered as supportive evidence for the view expressed by Hendry and Whiting (1972). They concluded that a closed system existed in physical education where traditional values may be perpetuated by teachers pursuing a 'narrow' curriculum and colleges of education failing to emphasise the subject's educative functions.

Secondly, pupils outlined a possible counselling role for the physical education teacher both in the field of personal relations and careers guidance. On the other hand, they were aware, to some extent, of the neglect of less able pupils, and ideally hoped for consideration for all pupils.

Reasons for this apparent anomaly can be advanced. Musgrove and Taylor (1969) have discussed the strong reinforcing influence of colleagues on the teacher's role, and the present study showed that, in agreement with colleagues, physical education teachers perceived an orientation towards the school's sporting success as valuable to their prestige. Hoyle (1969) has asserted that élitist values pervade the whole educational system, and so this situation is obviously not unique to physical education, only achieved in a different sphere of endeavour from academic subjects. While physical education teachers and colleagues shared the view that the physical education teacher was a competent, socially-oriented teacher who could establish good relations with pupils, there may have been staffroom pressures steering the physical education teacher towards élitism. As Hoyle (1969) has written, 'It is clear that the teacher who seeks to "educate for citizenship" . . . would be deviating considerably from the role expectations of his colleagues'.

Lastly, physical education teachers felt strongly a sense of low status, subject marginality, and the need for conformity, despite clear indications of impartiality by head teachers. Cohen (1970) has demonstrated that head teachers see themselves under obligation to show impartiality to all their staff, and this was reflected in the present findings. An alternative interpretation might suggest that social desirability of response was operative here.

The fact that headteachers were at odds with physical education teachers over the claims for greater efficiency or dedication on the part of these specialist teachers, or the way in which their subject was compared with other subjects, gives an interesting insight into the differing social perspectives which seem to operate within schools, and which can influence attitudes and subsequent social exchanges.

Pupils, colleagues and head teachers appeared to hold certain important, and different, expectations *vis-à-vis* the physical education teacher (and the majority of physical education teachers agreed with all of these views). Pupils wanted a counsellor, head teachers a disciplinarian, and colleagues desired competitive success for the sake of prestige.

Such differing orientations may well explain why a non-examinable subject like physical education is ill-defined in the academic hierarchy of the school. It is difficult to fulfil such diverse obligations, and the result is usually alignment with the more powerful role definers (Cain, 1968).

Thus, it seems reasonable to hypothesise that there is an interaction between the low status of the teacher of physical education and the personality qualities with which he is attributed—extraversion, aggressiveness and poor social perceptiveness (Hendry and Whiting, 1972). The product of this interaction may be a counter drive for achievement, reinforced by colleagues' adherence to and support for élitist values, where intense concentration on the competitive performance of the gifted few can occur. As a result, pupils may lose much of the educational benefits of the subject, and physical education teachers may miss the opportunity of a useful counselling role in the school.

References

APPLEBY, S. F. (1967). 'The role of the male teacher of physical education as perceived by men students in a college of education'. Unpublished diploma dissertation, University of Manchester.

BIGLIN, A. (1968). 'Some sources of conflict in the role of the physical education specialist', *The Leaflet*, 69, 10, 89–90.

BROWN, J. (1956). 'A modified significance test for the difference between two observed proportions', *Occupat. Psychol.*, 30, 169–74.

CAIN, M. E. (1968). 'Some suggested developments for role and reference group analysis', *Brit. J. Sociol.*, 19, 2, 191–205.

CARTER, J. E. L. (1964). 'The physiques of male physical education teachers in training', *Phys. Educ.*, 56, 66–76.

COHEN, L. (1967). 'The teacher's role and liaison between school and neighbourhood', in M. CRAFT, J. RAYNOR and L. COHEN (Eds.) *Linking Home and School*. London: Longmans.

COHEN, L. (1970). 'Conceptions of headteachers concerning their role'. PhD thesis, University of Keele.

HARGREAVES, A. (1968). Physical education in the BEd degree. Unpublished diploma dissertation, University of Keele.

HEATON, J. (1969). 'The role of the male physical education teacher working with boys in secondary schools'. Unpublished diploma dissertation, University of Birmingham.

HENDRY, L. B. and WHITING, H. T. A. (1972). 'General course and specialist physical education student characteristics', *Educ. Res.*, **14**, 2, 152–6.

HOYLE, E. (1969). *The Role of the Teacher*. London: Routledge & Kegan Paul.

LEE, S. F. (1968). 'The role of the male teacher of physical education as perceived by men students'. Unpublished diploma dissertation, University of Manchester.

MUSGROVE, F. and TAYLOR, P. H. (1969). *Society and the Teacher's Role*. London: Routledge & Kegan Paul.

REID, M. (1972). 'Comprehensive integration outside the classroom', *Educ. Res.*, **14**, 2, 128–34.

SCOTLAND, J. (1964). 'The physical education teacher', *Scot. Bull. Phys. Educ.*, 1, 2.

WARD, E. (1968). 'The use of tests in the selection of students for specialist physical education colleges for women'. MEd thesis, University of Manchester.

WESTWOOD, L. F. (1967). 'The role of the teacher'. *Educ. Res.*, **9**, 2, 122–34 and **10**, 1, 21–37.

The Impact of Teacher Life Changes and Performance on Student Dropouts

Elihu Carranza, Ph.D., Provost, Evergreen Valley College, San Jose, California

Summary

This study constitutes an assessment of the relation between the number of students leaving school (dropouts) and the magnitude of teacher life changes in conjunction with selected teacher performance variables in terms of pair-wise and multiple correlation procedures.

Findings support the thesis that the selected teacher performance variables taken in conjunction with teacher life changes and the number of student dropouts are significantly associated.

Introduction

There is now a growing body of evidence that supports the claim that acceleration of change beyond the limits the human organism can absorb results in both physical and psychological distress, the effects of which may lead to illness or breakdown (Brown and Birley, 1968; Cassel and Tyroler, 1961; Cuadill, 1958; Hinkle and Wolf, 1958; Graham and Stevenson, 1963; Rahe *et. al.*, 1967, 1968, 1970). In a recent study an extension of this claim to determine the effects of life changes (such as marriage, divorce, death of spouse, sex difficulties, etc.) on teacher performance was investigated (Carranza, 1972). The study adduced evidence that life changes and teacher performance variables (such as teacher absence-illness, absence-frequency, absence-duration, requests for transfer, job moves, etc.) were significantly correlated. While the study confirmed a number of statements concerning the relation between life changes and teacher performance variables, the statistical methods used did not reveal a significant correlation ($<\cdot 05$) between teacher life changes and student dropouts. It was left as a project for further research to determine whether the teacher performance variable, student dropouts, was significantly associated with *both* the total life changes experienced by the teacher *and* the remaining set of teacher performance variables (or any subset of teacher performance variables).

For purposes of this study the term *dropout* denotes a student eighteen years of age or younger who in effect was truant one or more times during some portion of his school career. The duration of truancy, i.e., the length of the student absence without permission is closely monitored by the school district since funding depends in part on total teacher-student contact hours. A truant student is designated as a dropout when it becomes clear, for example, that he has left his home or the board and care of his guardian and after a reasonable period of time an investigation determines that the student is absent without permission and intends to remain so.

Purpose of the study

The actualization of the student's potential in the classroom has come to depend in large measure upon the success or failure of his relations with his teacher. Such success or failure, however, appears to be a function of many subtle factors which interact to influence the behaviour that in increasingly many instances terminates in student truancy (Cervantes, 1965). This study sought to identify some of those subtle factors which could prompt students illegally to terminate their school careers. An assessment of the relation between the number of student dropouts and the life changes experienced by the teacher in conjunction with certain selected teacher performance variables was undertaken by means of Pearson's correlation and multiple correlation coefficients.

It is important to keep in mind that this study sought to determine the strength of the postulated relationship. Cause-effect determinations fell outside the scope of the study.

Procedures

Life changes experienced by the teacher were measured by the *Schedule of Recent Experience* (SRE) and the *Social Readjustment Rating Scale* (SRRS) (see Appendix) developed by Holmes and Rahe (1967 and Holmes and Masuda, 1970).

The SRE comprises forty-two unique life event items divided into two categories: (1) those indicative of the life style of the individual, and (2) those indicative of occurrences involving the individual. These life change events cover a broad spectrum of ordinary social and interpersonal transactions ranging from matters of occupation, residence, community, family and marriage to matters of religion, personal habits, and health. The life events described in both sections share a common characteristic in that an occurrence of a life event usually is associated with some coping or adaptive behaviour on the part of the individual concerned.

The SRRS is the product of a method for quantifying the amount of change in life adjustment required by the life events described in each of the forty-two items of the SRE. The SRRS assigns a certain magnitude denoted by a number (Life Change Unit or LCU) to each of the life change items according to the amount, severity and duration of adjustment each item (in comparison to other items) is perceived to require. Thus, a subject's life changes measured by the SRE are scaled numerically according to the amount of life pattern change estimated for an individual and the degree of readjustment required in coping with each change (Holmes and Rahe, 1967; Holmes and Masuda, 1970). The scale ranges from 11 LCU (minor violations of the law) to 100 LCU (death of a spouse). The position on the scale of any life event, then, is indicated by the number of LCU estimated to apply to the life change event in question. The result is quantification of life changes for a specified period of time in terms of LCU scores.

According to Holmes and Masuda, the method for quantifying life changes was derived from psychophysics (Holmes and Rahe, 1967; Holmes and Masuda, 1970; Masuda and Holmes, 1967a, Psychophysics may be defined as the study of the psychological perceptions of the quality, quantity, magnitude, and the intensity of physical phenomena. Stevens (1951) defines psychophysics as the science of the response of organisms to stimulating configurations. According to Stevens, the response can be viewed as an indicator of an attribute of an individual—an attribute which varies with the stimulus and, perhaps more important, is relatively invariant from person to person.

In developing the SRRS, Holmes and Rahe postulated that participants' psychological capacity for making quantitative judgements about psychosocial and psychophysical phenomena provide a reliable basis for the quantification of life change process. Data generated in studies by Holmes and Rahe and others support the validity of the assumption (Holmes and Rahe, 1967; Holmes and Masuda, 1970; Masuda and Holmes, 1967a, 1967b; Casey *et al.*, 1967). The results of these studies demonstrate a high degree of consensus suggesting agreement among individuals and between groups which transcend or cut across differences in marital status, age, sex, education, social class, generation American, religion and race.

Data for the selected variables of absence-illness, absence-frequency, absence-duration, requests for transfer, number of job moves within the district, units earned beyond the Bachelor of Arts degree, number of students, number of classes, percentage of student A's and B's, and percentage of student D's and F's (Excellent, Good, Poor, and Failing marks, respectively) on teacher performance were collected from various sources through the Office of Research of the school district selected for this study.

The population under investigation consisted of high school teachers from three senior high schools in the Lansing School District, Lansing, Michigan with teacher contracts for the years 1970-71 and 1971-72. A sample of N=110 out of 152 who responded to a letter soliciting teacher participation provided the basis for this study.

Description of selected variables

1. *Total LCU Score* is a number denoting the total amount of life changes as measured by the SRE which the high school teacher experienced during the past three years from December, 1968 to December, 1971.

2. *Teacher Absenteeism-Illness* is comprised of total number of days the teacher was absent due to reported illness or injury during the time period beginning September, 1970 to December, 1971.

3. *Teacher Absenteeism-Frequency* is comprised of the total number of occasions the teacher was absent during the time period beginning September, 1970 to December, 1971.

4. *Teacher Absenteeism-Duration* is comprised of the total number of days the teacher was absent during the time period beginning September, 1970 to December, 1971.

5. *Transfer Requests* refers to all requests either by the teacher or an administrator for the reassignment or transfer of the teacher within the district or within the school during the time period beginning September, 1970 to December, 1971.

6. *Number of Job Moves* refers to all distinct job moves by the teacher within the district or within the schools during the time period beginning September, 1970 to December, 1971.

7. *Units Earned* refers to the number of units or credits earned by the teacher beyond the Bachelor of Arts degree in terms of one of the following categories: (1) no units beyond the Bachelor of Arts degree, (2) Bachelor of Arts degree plus 23 units, (3) Master of Arts degree, or (4) Master of Arts degree plus 45 or more additional units during the time period beginning September, 1970 to December, 1971.

8. *Number of Students* refers to the total number of students assigned to the teacher for the Fall term beginning September, 1971 to January 28, 1972.

9. *Number of Classes* refers to the total number of classes taught by the teacher during the Fall term beginning September, 1971 to January 28, 1972.

10. & 11. *Percentage of A's and B's* and *Percentage of D's and F's* refers to the percentages of high and low student grades, respectively, distributed to students by teachers during the time period beginning September, 1971 to January 28, 1972. An 'A' grade denotes *Superior* or *Excellent* (90 per cent to 100 per cent); a 'B' grade denotes *Good* (80 per cent to 89 per cent); a 'C' grade denotes *Average* (70 per cent to 79 per cent); a 'D' grade denotes *Poor* (60 per cent to 69 per cent); and an 'F' grade denotes *Failure* (59 per cent and below).

12. *Student Dropouts* is composed of the total number of student dropouts who as former students attended one or more of the classes of one or more of the teachers and described (by the Child Accounting section of the school district) as persons under eighteen years of age who were truant or otherwise absent from school for a period of time without proper authority on one or more occasions during the time period beginning September, 1970 to December, 1971. It should be noted that dropouts generally do not graduate.

13. *Teacher Performance* is a construct of the teacher performance variables described above.

Statistical methods

The hypothesis under investigation states that there is no significant correlation between high school teacher total LCU scores and selected teacher performance variables of (1) absence-illness, (2) absence-frequency, (3) absence-duration, (4) transfer requests, (5) number of job moves, (6) units earned beyond the Bachelor of Arts degree, (7) number of students, (8) number of classes, (9) percentage of student A's and B's, and (10) percentage of student D's and F's, when simultaneously included in a multiple regression equation, and student dropouts.

On the assumption that some of the correlation coefficients were significantly different from zero, a step-wise deletion of variables was used (a) to determine whether the deletion of variables produced a significant sub-set, i.e., a sub-set of teacher performance variables which contributed significantly to the overall prediction, and (b) to achieve the most parsimonious, yet precise, solution possible. The step-wise deletion of variables procedure deletes those variables in the original equation which did not contribute significantly to the overall prediction of a multiple regression equation.

In sum, multiple correlation analysis or multiple regression provides an analysis of the relations among a single criterion variable (number of student dropouts) and two or more predictor variables (teacher performance variables) where total LCU scores of high school teachers are automatically retained in the regression equations proceding until the minimum significance criterion was satisfied.

The ·05 level of significance was selected as

Table 1: Correlation matrix of variables

	Total LCU	Absence–Illness	Absence–Frequency	Absence–Duration	Transfer Requests	Job Moves	Units Earned	Number of Students	Number of Classes	%A's and B's	% D's and F's	Student Dropouts
1. Total LCU		·21					−·22					
2. Absence–Ill.			·55	·76			−·28					
3. Absence–Freq.				·85			−·24					
4. Absence–Dur.							−·25					
5. Transfer Req.						·64						−·19
6. Job Moves										·24	−·21	−·19
7. Units Earned								·33				
8. No. Students									·25			·27
9. No. Classes												·24
10. % A's and B's											−·85	−·19
11. % D's and F's												·23
12. Student Dropouts												

$R \geqslant \pm \cdot 187, p \leqslant \cdot 05$

the critical region of rejection.

Findings

Table 1 depicts the pair-wise correlation coefficients between the variables. The correlation coefficients which had a value of $R \geqslant \pm \cdot 187$ were significant at $p \leqslant \cdot 05$ level.

The more relevant correlation coefficients are depicted in the column under the variable student dropout. The highest correlations were between student dropouts and (1) number of students (·27), (2) number of classes (·24), and (3) percentage of student D's and F's (·23). The remaining significant correlations between student dropouts and (4) transfer requests, (5) teacher job moves, and (6) percentage of student A's and B's were negative and had the same value (—·19).

In sum, these findings indicate that the magnitude of student dropouts are positively correlated with the number of students, number of classes, and the percentage of student D's and F's distributed by the teacher. The negative simple correlation coefficients between student dropouts and teacher transfer requests, teacher job moves, and percentage of student A's and B's indicate that the greater the number of student dropouts the less the percentage of student A's and B's, the less job moves experienced by the teacher, and the less the number of requests for transfer or reassignment within the school or district.

Table 2: Relationship between student dropouts and TPV including total LCU scores

Criterion Variable	Number of TPV Deleted	R	DF	F	P
Student Dropouts	None	·44	11/98	2·17	·022

Table 2 presents the multiple correlation coefficient, F-ratio, degrees of freedom, and the level of significance for the joint relationship between student dropouts as a criterion variable (dependent variable) and the significant teacher performance variables (TPV) used as predictors (independent variables) including the total LCU scores prior to deletion of variables.

Although these data exceed the ·05 alpha level of significance, they do not represent the most parsimonious solution possible. These data indicate that the set of TPV together with the total LCU scores of teachers, significantly predict student dropouts.

Table 3 presents the multiple correlation coefficient, F-ratio, degrees of freedom, and the level of significance for the joint relationship of student dropouts as a criterion variable and the significant, remaining (after deletion) TPV which together with total LCU scores were used as

Table 3: Relationship between student dropouts and the most parsimonious sub-set of TPV including total LCU scores

Criterion Variable	Number of TPV Deleted	R	DF	F	P
Student Dropouts	7	·41	4/105	5·18	·001

predictors. The most parsimonious sub-set of TPV which together with total LCU scores significantly predict student dropouts were (1) number of students, (2) number of classes, and (3) per cent student D's and F's. Variables were deleted according to the statistical significance of their contribution to the prediction of the criterion.

The deleted teacher performance variables were (1) percentage of student A's and B's, (2) absence-frequency, (3) absence-illness, (4) units earned, (5) job moves, (6) absence-duration, and (7) requests for transfer—in that order. The deletion of seven TPV as predictors in the regression equation with student dropouts as criterion, produced an increase in the level of significance from ·022 (R=·44) to ·001 (R=·41).

The most parsimonious, yet precise, solution possible can be stated as a prediction equation formed from the regression coefficients of the remaining (after deletion) significant TPV.

Student Dropouts= —5·12+·002 (Total LCU) +·077 (Number of Students)+1·88 (Number of Classes)+·25 (Percent Student D's and F's).

Summary and discussion

The purpose of this study, as previously noted, was to determine the impact of teacher life changes and teacher performance, as defined, on student dropouts by means of simple and multiple correlations between student dropouts and selected TPV and total LCU scores of high school teachers.

Generally, significant, positive, simple correlation coefficients between student dropouts and (1) number of students, (2) number of classes, and (3) percentage of student D's and F's, as well as the significant, negative, simple correlation coefficients between student dropouts and (4) transfer requests, (5) job moves, and (6) percentage of student A's and B's indicate that some obvious as well as more subtle factors are operant in student dropout behaviour.

Thus teacher performance variables, and generally, those aspects of teacher performance variables considered less desirable by many teachers (more students, more classes, more D's and F's, and less A's and B's) are significantly associated with the number of student dropouts.

In addition, total LCU scores are negatively correlated with units earned (—·22) and these are positively correlated with number of students assigned to the teacher (·33). The greater the number of students assigned to a teacher the greater the number of dropouts (·27). Thus, life changes experienced by the teacher probably constitute an important variable worthy of careful study in the assessment of student dropout rates.

Further, since the multiple correlation coefficients of R=·44 (student dropouts and the set of ten TPV plus total LCU scores) and R=·41 (student dropouts and a sub-set of three TPV plus total LCU scores) were significant at less than the ·05 alpha level, the null hypothesis of no significant correlation was rejected. The deletion of seven TPV resulted in greater significance and produced the most parsimonious, yet precise, solution possible. Such a solution retained the TPV of (1) number of students, (2) number of classes, and (3) percentage of student D's and F's (together with total LCU scores) in the prediction equation as predictors of the number of student dropouts.

In sum, this study adduces evidence to support the thesis that the selected teacher performance variables plus LCU scores and the number of student dropouts are significantly correlated. Moreover, the evidence indicates that mobility of the teacher may be an important factor in high student dropout rates. Finally, the number of students and the number of classes assigned to the high school teacher and the percent D's and F's distributed to students significantly predict or are otherwise significantly related to student dropouts.

Appendix: The Social Readjustment Rating Scale*

No.	*Life Event*	*Mean Value*
1.	Death of spouse	100
2.	Divorce	73
3.	Marital separation	65

---

*See Holmes, T. H., and Rahe, R. H. (1967). 'The social readjustment rating scale', *J. Psychosomatic Res.*, **11**, 213–18, for complete wording of the items.

No.	Life Event	Mean Value
4.	Jail term	63
5.	Death of close family member	63
6.	Personal injury or illness	53
7.	Marriage	50
8.	Fired at work	47
9.	Marital reconciliation	45
10.	Retirement	45
11.	Change in health of family member	44
12.	Pregnancy	40
13.	Sex difficulties	39
14.	Gain of new family member	39
15.	Business readjustment	39
16.	Change in financial state	38
17.	Death of close friend	37
18.	Change to different line of work	36
19.	Change in number of arguments with spouse	35
20.	Mortgage over $10,000	31
21.	Foreclosure of mortgage or loan	30
22.	Change in responsibilities at work	29
23.	Son or daughter leaving home	29
24.	Trouble with in-laws	29
25.	Outstanding personal achievement	28
26.	Wife begin or stop work	26
27.	Begin or end school	26
28.	Change in living conditions	25
29.	Revision of personal habits	24
30.	Trouble with boss	23
31.	Change in work hours or conditions	20
32.	Change in residence	20
33.	Change in schools	20
34.	Change in recreation	19
35.	Change in church activities	19
36.	Change in social activities	18
37.	Mortgage or loan less than $10,000	17
38.	Change in sleeping habits	16
39.	Change in number of family get-togethers	15
40.	Change in eating habits	15
41.	Vacation	13
42.	Christmas	12
43.	Minor violations of the law	11

References

Brown, G. W., and Birley, J. L. T. (1968). 'Crises and life changes and the onset of schizophrenia', *J. of Health and Soc. Beh.* **9**, 203.

Carranza, E. (1972). 'A study of the impact of life changes on high school teacher performance in the Lansing school district as measured by the Holmes and Rahe schedule of recent experience'. Unpublished doctoral thesis, Michigan State University.

Casey, R. L., Masuda, M., and Holmes, T. H. (1967). 'Quantitative study of recall of life events', *J. Psychosomatic Res.*, **11**, 239.

Cassel, J., and Tyroler, H. (1961). 'Epidemiological studies of culture change', *Archives of Environment and Health*, **3**, 25.

Caudill, W. (1958). *Effects of Social and Cultural Systems in Reactions to Stress*. (Pamphlet 14.) New York: Social Science Research Council.

Cervantes, L. F. (1965). *The Dropout, Causes and Cures*. Ann Arbor: The University of Michigan Press.

Graham, D. T., and Stevenson, I. (1963). 'Disease as response to life stress', in Lief, H. I., V.F. and N.R. (Eds.). *The Psychological Basis of Medical Practice*. New York: Harper and Row.

Hinkle, L. E., Jr. and Wolf, H. G. (1958). 'Ecological investigations of the relationship between illness, life experiences and the social environment', *Annals of Internal Medicine*, **49**, 1373.

Hinkle, L. E. Jr., Christensen, W. N., Kane, F. D., Ostfeld, A., Thetford, W. N. and Wolf, H. G. (1958). 'An investigation of the relation between life experience, personality characteristics and general susceptibility to illness', *Psychosomatic Medicine*, **20**, 278.

Holmes, T. H. and Masuda, M. (1970). 'Life changes and illness susceptability', from *Separation and Depression*. American Association for the Advancement of Science, 161–186.

Holmes, T. H. and Rahe, R. H. (1967). 'The social readjustment rating scale', *J. Psychosomatic Res.*, **11**, 213.

Masuda, M. and Holmes, T. H. (1967a). 'Magnitude estimations of social readjustments', *J. Psychosomatic Res.*, **11**, 224.

Masuda, M. and Holmes, T. H. (1967b). 'The social readjustment rating scale: a cross-cultural study of Japanese and Americans', *J. Psychosomatic Res.*, **11**, 227.

Rahe, R. H. (1964). 'Social stress and illness onset'. *J. Psychosomatic Res.*, **8**, 35.

Rahe, R. H. (1968). 'Life change patterns surrounding illness experience', *J. Psychosomatic Res.*, **11**, 341.

Rahe, R. H., Mahan, J. L. Jr. and Ransom, J. A. (1970). 'Prediction of near-future health change from subjects' preceding life changes', *J. Psychosomatic Res.*, **14**, 401.

Rahe, R. H., McKean, J. D. Jr. and Ransom, J. A. (1967). 'A longitudinal study of life-change and illness patterns', *J. Psychosomatic Res.*, **10**, 355.

Stevens, S. S. (Ed.) (1951). 'Mathematics, measurement, and psycho-physics', *Handbook of Experimental Psychology*. New York: John Wiley and Sons, Inc.

Adjustment to School

Barbara Thompson, Staff Tutor in Social Science, Open University

Summary

This article reports the development of an instrument to measure initial adjustment to school, and its use to explore the differences between children with and without nursery experience. Those with nursery experience proved to have a better grounding for cognitive work in infant school, but there was no significant difference on emotional or social criteria of adjustment to school.

1. Rationale

'Readiness is necessary for success' (Austin and Lafferty, 1968). The child who is ready for school and hence makes a satisfactory initial adjustment to it is more likely to be successful in the rest of his educational career than the child who, because he is not ready, finds difficulty in coping with the school situation.

Studies in America by Ilg and Ames (1964) and in Sweden by Johansson (1965) have found substantial correlations between assessments of school readiness and later performance. If one assumes that these findings can be generalised to the United Kingdom, it becomes important to explore what early experiences are likely to be related to satisfactory adjustment.

In order to do this there must be some criteria of adjustment. The instruments used in the studies mentioned above are clearly not satisfactory as not only do they derive from different cultural backgrounds, but are intended for children starting school at a later age than in the United Kingdom.

The current study describes the development of an instrument to measure initial adjustment to school which can be used by teachers to assess children in their first term.

2. The effect of nursery education on adjustment to school

One important use to which such an instrument can be put, is to examine the relative success of children with and without previous experience in nursery schools and classes (in coping with school). Despite the constantly expressed belief in the value of pre-school education, surprisingly little hard evidence exists for its effectiveness, and the results that are available are by no means unequivocal.[1]

Two main areas have been investigated, that of linguistic and other intellectual performance, and that of social and emotional maturity.

(a) *Linguistic and other intellectual performance*

Douglas and Ross (1964) found that within each social class former nursery children obtained higher than average scores on a combined measure of achievement, at the age of eight, although none of the differences reached significance. Harrold and Temple (1959-60) found children with nursery experience superior on vocabulary and expression measured from a free speech sample to those with no nursery experience, while Cohen and Bagshaw (1973) found significant improvement in scores on the Communication Scale of the Progress Assessment Chart (Gunzberg, 1969) after a year's nursery experience. Widlake (1971), however, found no superiority for his nursery group on the English Picture Vocabulary Test; probably due to the fact that this test is concerned with receptive, not expressive, vocabulary which would seem less likely to be affected by nursery experience.

---

[1] No reference is made here to studies carried out in the USA, or to the effects of special Head Start type programmes.

Stukat (1963), in a study carried out in Sweden, found children with pre-school (kindergarten) experience more interested in manual-constructive activities, and superior on general knowledge, and coherence and correctness of language use than those without pre-school experience. In a second Swedish study, however, Johansson (1965) found that although the means for ex-kindergarten children were higher for those without pre-school experience on all his measures of readiness and school performance—except reaction to orders—none of these differences proved statistically significant.

(b) *Social and emotional maturity*

In the area of social-skills, Widlake, using the Vineland Social Maturity Scale, (Doll, 1965), found his nursery group superior, particularly as regards self-help, eating, dressing and locomotion; and Harrold and Temple found the ex-nursery children better than non-nursery in their (i) attitude to other children, (ii) self-confidence, (iii) response to environment, (iv) response to strange adults, and (v) ability to dress themselves. In the Swedish studies, Stukat's only significant findings were that ex-nursery children were better able to fend for themselves, and Johansson found they had a significantly better attitude to their classmates than children without pre-school experience. However, Cohen and Bagshaw found no improvement on the Self-help and Socialization scales of the PAC1, and when behaviour of the ex-nursery children from the Douglas and Ross study was assessed at 13 and 15, a higher proportion than expected were rated as highly maladjusted. Although the latter result might be explained in terms of the policy of nursery schools to take in children at risk, the limited superiority of the children in the two Swedish studies, and the total lack of it in Cohen and Bagshaw's, is less easy to explain.

Stukat claims that comparisons of this kind are often invalidated by the difficulty of finding equivalent groups. The studies quoted here have used a variety of methods to attain such equivalence. Stukat's own was probably the most thorough. A 'home group' supplied the controls for the experiment. These were children selected from the long waiting lists of unsuccessful applicants for places at the same kindergartens as the experimental children who had received pre-school training, and carefully matched with them on the basis of sex, age, father's occupation, area of domicile and number of siblings. The Widlake and Harrold and Temple studies also used matched pairs of children, although the matching was on a less exhaustive set of variables. Widlake admits that 'We were very much aware that conditions in the families differed considerably'. Douglas and Ross 'made allowances for the combined effects of maternal care, housing and family size, as well as social class' and Johansson, keeping intelligence constant, found his nursery and non-nursery groups differed on parental education but not on economic status, social class, sex or age. Finally, Cohen and Bagshaw used the children as their own controls, with a pre-test/post-test design.

It seems possible that the variety of ways in which comparison groups were chosen, together with the different ages at which testing took place (these ranged from eight in the Douglas and Ross study to four for half of Cohen and Bagshaw's sample), accounts for the differences in result.

The Adjustment to School Scale, discussed in this article, has initially been used to rate all new entrants to a number of schools. The data have then been analysed:

(a) to examine the structure of the instrument;

(b) to test the hypothesis that children who have had experience in nursery schools will adjust to school better than those who have not.

3. The development of the scale

It was decided to use a rating scale which would be completed by the child's teacher early in the second half of his first term at school. Despite the known limitations of rating scales, especially when completed by untrained raters, this appeared the only practical method of gathering information about children of this age.

In order to make the rating task relatively simple, it was decided to have only four categories. A scale of this size, having no mid-point, was also felt to have the advantage of forcing the raters to discriminate amongst the 'average' group of children.

It seemed likely that three main areas of behaviour would generally be accepted as indicative of how well a child was coping with school.

(a) *His personal emotional adjustment*—that is to say, the extent to which he is independent, and self-reliant, and can accept criticism of his work and behaviour with good grace.

(b) *His social adjustment*—whether he relates satisfactorily to his teachers and his peers, and can cope maturely with unfamiliar adults.

(c) *His attitudes and behaviour in response to the intellectual demands of the school*—the extent to which he is interested in his work and surroundings; the use he makes of opportunities offered to him.

The first two of these areas have traditionally been of great concern to teachers in nursery schools. The recent Schools Council Working Paper (Taylor *et al.*, 1972) found 'Socio-emotional development' was ranked first amongst five agreed purposes of nursery education. The purposes ranked second and third in the same study were 'Intellectual development' and 'Creation of an effective transition from home to school'. These to a considerable extent cover the third area discussed above. However, the whole question of cognitive development seems to induce an ambivalent attitude. On the one hand nursery staff realize that they are preparing children to take their places in the school system, whilst on the other they seem afraid of 'stealing their childhood', of hastening the children's progress towards the adult world of work. Hence, cognitive skills may be left to be acquired indirectly and incidentally, and evidence from some of the American studies (Bereiter, 1972; Weikart, 1972) would seem to suggest that this is most unlikely to be successful. For the infant school teacher, on the other hand, faced with a class of thirty five-year-olds, it is as important for the children to possess the skills and attitudes which are needed to cope with the intellectual aspects of school as for them to be mature enough socially and emotionally. Indeed if either aspect is missing, it causes additional problems for the teacher. Husen (1962) defines school readiness in terms of 'the demands made by the teacher on the pupils' ability to join a group at work, to treat classmates with consideration, to satisfy the demands made by elementary instruction in reading, writing and mathematics' (quoted in Johansson, 1965, p.26).

After discussion with a number of nursery and infant school teachers, an instrument was prepared consisting of twenty items, each to be rated on a four-point scale. These it was hoped represented those aspects of the child's behaviour which the nursery school considered of importance as well as those expected of the child by teachers in the infant school.

The scales covered the following areas:

(1)	Settling in school	(S)
(2)	Co-operation with other children	(S)
(3a)	The child's relationship with the teacher (i)	(S)
(3b)	The child's relationship with the teacher (ii)	(S)
(4)	Attention seeking	(P)
(5)	Need for teacher's approval	(P)
(6)	Acceptance of criticism and blame	(P)
(7)	Reaction to strange adults	(S)
(8)	Level of concentration	(C)
(9)	Use of play materials	(C)
(10)	Attitude to work	(C)
(11)	Creativity and originality	(C)
(12)	Interest in surroundings	(C)
(13)	Attitudes towards new activities	(C)
(14)	Self-reliance	(P)
(15)	Verbalisation of wants and needs	(Cv)
(16)	Verbalising abilities in school work	(Cv)
(17)	Following instructions	(Cv)
(18)	Behaviour at story time	(Cv)
(19)	Ability to cope with personal needs	(P)

(The letters in brackets indicate the areas each item was thought to exemplify. (S) Social, (P) Personal/emotional, (C) General cognitive, (Cv) Cognitive/Verbal.)

4. The structure of the instrument

(a) *The pilot investigation*

(i) *Data collection.* In order to explore the factorial structure of the instrument, data were collected on a sample consisting of all pupils entering the reception class or its equivalent in ten inner metropolitan schools, in January 1971. This is a term when intake tends to be low, as most heads like to take as many children as possible into school at the beginning of the academic year, even if they are still below five years of age. The total sample was therefore somewhat smaller than had been hoped. However, forms were completed for 137 children.

Of the children for whom forms were returned, 83 had spent some time in a nursery school or class. A number of others had attended play groups. As no information was available as to how the latter were staffed, or what opportunities they provided, children with playgroup experience were included with those entering school direct from home. It would be expected that all playgroups would provide some social training, and therefore differences between the two groups of children in this area would be somewhat reduced.

No attempt was made to collect full details of socio-economic class or other background variables. However, from information obtained from head teachers, it was apparent that the majority of the sample would be classified as working class. Most of the schools had a fair proportion of children from overseas.

(ii) *Procedure*. Teachers were asked to complete a form for each child who had started school that term, by ringing the number against the statement in each scale which they thought most applicable to him. No space was provided on the form for information about pre-school experience in an attempt to guard against any bias induced by having the child's pre-school history before the teachers as they worked. It was, of course, impossible to guard totally against bias, as many teachers would have been well aware of whether or not the child had attended nursery.

The heads were asked to enclose a list with the completed scripts, giving name, date of birth and nursery experience, if any. Any additional information was welcomed.

(iii) *Factor analysis*. Ratings on the twenty scales were correlated and a factor analysis carried out. Three factors were extracted which together accounted for just over 57 per cent of the total variance. Table 1 shows the loadings for each item on the three factors after rotation by the Varimax procedure. Examination of these loadings confirms that the scales fall into three groups which approximate very closely to those predicted when the instrument was constructed. Factor 1, which accounts for by far the largest amount of the variance (39 per cent) consists of items relating to '*attitudes and behaviour in response to the intellectual demands of the school*'. It is perhaps most appropriately labelled 'Learning readiness'. Factor 2, accounting for 11 per cent of the variance, consists of items relating to '*competence in social relations*' although Scale 2, 'co-operation with other children', loads more heavily on Factor 1. An additional item, 13, 'Attitude towards new activities' is included here, perhaps indicating the social component of new events in school. Finally, Factor 3, which accounts for only 7 per cent of the variance, relates to '*Personal Maturity/Independence*' but has lost to Factor 1, Item 19, 'Ability to cope with personal needs'. This is clearly an important aspect of the child's readiness for school. It will be noted that Item 14, 'Self-reliance', has been included here despite the fact that its loadings are marginally higher on both the other factors. As the differences in loadings were so small, the item was allocated to the factor which appeared psychologically most meaningful.

(iv) *Reliability*. Reliability, in the form of internal consistency, was measured by calculating alpha co-efficients by the 'short cut' formula (McKennell, 1970) for the clusters of scales

Table 1: Loadings of items on the three factors derived from the pilot data

Item No.		*Factor 1*	*Factor 2*	*Factor 3*
17	Following instructions	*–0·79*	–0·03	–0·11
10	Attitude to work	*–0·75*	–0·29	–0·05
16	Verbalising ability in schoolwork	*–0·74*	–0·24	0·17
18	Behaviour at story time	*–0·72*	–0·01	–0·20
15	Verbalisation of wants and needs	*–0·72*	–0·19	0·23
8	Level of concentration	*–0·68*	–0·14	–0·43
9	Use of play materials	*–0·65*	–0·37	–0·12
11	Creativity and originality	*–0·61*	–0·45	–0·02
12	Interest in surroundings	*–0·61*	–0·57	0·12
2	Co-operation with other children	*–0·49*	–0·33	–0·28
19	Ability to cope with personal needs	*–0·43*	–0·20	–0·08
3a	Child's relation with teacher (1)	–0·26	*–0·83*	0·10
3b	Child's relation with teacher (2)	0·02	*0·79*	–0·13
7	Reaction to strange adults	–0·30	*–0·71*	0·10
1	Settling in school	–0·15	*–0·67*	–0·24
13	Attitude towards new activities	–0·44	*–0·57*	–0·17
4	Attention seeking	–0·05	–0·31	*0·72*
5	Need for teacher's approval	–0·07	0·14	*0·69*
6	Acceptance of criticism and blame	–0·25	0·00	*–0·54*
14	Self-reliance	–0·49	0·47	*–0·41*

represented by the three factors.

The results, shown in Table 2, indicate that while the values obtained for Factors 1 and 2 are satisfactory, that for Factor 3 is less so.

Table 2: Internal consistency co-efficients for factors obtained from pilot data

	Factor	*No. of items*	*Alpha co-efficient*
1	School readiness	11	0·90
2	Social competence	5	0·77
3	Personal maturity	4	0·56

(b) *The Slough sample*

(i) *Data collection.* As part of the NFER pre-school project, 353 children entering eleven Slough infant schools in the terms beginning April 1971, September 1971, and January 1972, were rated on the Adjustment to School Scale in the second half of their first term at school. At the same time, the Boehm Test of Basic Concepts was administered. Of these children, 154 had attended a nursery school or class, and 199 had entered school direct from home.

Once again, information about pre-school experience was collected on a separate sheet after the teachers had rated the children, to avoid obvious bias.

(ii) *Factor analysis.* The responses to the twenty rating scales were intercorrelated, and the resulting matrix factor-analysed. On this occasion four factors were extracted as it seemed likely that the first factor obtained from the pilot data could be subdivided. Just over 64 per cent of the total variance was accounted for, the orthogonal factors accounting for 40 per cent, 11 per cent, 8 per cent and 5 per cent respectively. Loadings of the individual items on all four factors are shown in Table 3. Once again the first factor can be best labelled as 'Learning readiness', the second as 'Competence in social relations', and the third 'Personal maturity', although the self-reliance item, number 14, is now found to relate clearly to Factor 1. The new fourth factor is a verbal ability factor, comprising scales relating to 'verbalisation of wants and needs', 'verbalising ability in school work', and 'ability to cope with personal needs'.

(iii) *Reliability.* Alpha coefficients of internal consistency are shown in Table 4 below.

Table 4: Internal consistency co-efficients for factors obtained from Slough data

	Factor	*No. of Items*	*Alpha Co-efficient*
1	School readiness	10	0·92
2	Competence in social relations	4	0·72
3	Personal maturity	3	0·66
4	Verbal ability	3	0·63

Table 3: Loadings of items on four factors extracted from Slough data

Item No.		*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*
8	Level of concentration	*−0·82*	−0·03	−0·24	−0·08
10	Attitude to work	*−0·80*	−0·01	−0·13	0·09
9	Use of play materials	*−0·79*	−0·11	−0·18	0·13
11	Creativity and originality	*−0·72*	−0·24	0·05	0·19
12	Interest in surroundings	*−0·69*	−0·35	0·13	0·21
17	Following instructions	*−0·68*	−0·07	−0·05	0·41
18	Behaviour at story time	*−0·62*	−0·06	−0·29	0·14
13	Attitude towards new activities	*−0·61*	−0·42	−0·09	0·27
14	Self-reliance	*−0·59*	−0·26	−0·26	0·30
2	Co-operation with other children	*−0·53*	−0·37	−0·21	0·29
3a	Child's relationship with teacher (1)	−0·32	*−0·77*	0·20	0·16
3b	Child's relationship with teacher (2)	−0·14	*−0·75*	0·00	−0·11
1	Settling in school	−0·03	*−0·63*	−0·42	0·23
7	Reaction to strange adults	−0·04	*−0·58*	0·18	0·52
4	Attention seeking	0·09	−0·24	*0·82*	0·02
5	Need for teacher's approval	0·14	0·04	*0·76*	−0·10
6	Acceptance of criticism and blame	−0·37	−0·36	*−0·54*	−0·01
19	Ability to cope with personal needs	−0·19	0·00	−0·30	*0·70*
15	Verbalization of wants and needs	−0·49	−0·15	0·02	*0·70*
16	Verbalizing ability in schoolwork	−0·53	−0·14	0·07	*0·64*

On the whole the coefficients are more satisfactory for this set of analyses. This is particularly so in the case of Factor 3, which now has an alpha coefficient of 0·66 compared with 0·56 for the pilot data, even though it now has only three rather than four items.

(c) *Comparison of the factors obtained from the two analyses*

The loadings of the twenty items on the three factors obtained from the pilot sample and the four obtained from the Slough data were compared using the coefficient of congruence described by Harman (1960). Examination of the coefficients revealed that those comparing Pilot Factor 1 with Slough Factor 1 (0·96), and Pilot Factor 3 with Slough Factor 3 (0·94) were such as to suggest a very close resemblance between the factors. In addition the coefficient comparing Pilot Factor 1 and Slough Factor 4 (−0·75) was high enough to suggest a relation between these; a result not unexpected in that all three items comprising Slough Factor 4 formed part of the first factor in the pilot study.

The small size of the coefficient comparing the two second factors (only 0·57) is not as surprising as would appear at first sight. The coding of one of the component items, 3b, a scale which only applies to those rated 4 on Item 3a, was carried out differently on the two occasions. Whereas for the factor analysis of the pilot data, the responses were coded 1, 2, 3, 4 with 9 for 'not applicable' (the majority of responses fall in this category), by the time the Slough data were coded, these had been amended as follows:

Ratings 1, 2 and 4	coded 1
Rating 3	coded 2
Not applicable	coded 0

This change had the effect, among others, of reversing the signs of the loadings on the second occasion and hence markedly reducing the coefficient of congruence.

Overall, the results of the comparison suggest that there is considerable similarity between the factors obtained on the two samples and hence that the structure of the instrument is satisfactorily stable.

(d) *The relation between Adjustment to School and the Boehm Test of Basic Concepts*

As was stated above, at the same time as their teachers assessed their adjustment to school, the children took the Boehm Test of Basic Concepts (Boehm, 1970) which was administered by the project staff. This test is designed to measure children's mastery of the concepts considered necessary for achievement in the first years of school. It consists of two books each containing 25 pictorial multiple-choice items. The tester reads out the directions and the child is required to make a pencil cross on the item he selects. One major advantage of this test is that it can be administered to five-year-olds in groups of between six and ten children.

The scores of the two Boehm books were correlated with the ratings of the items of the Adjustment to School scale. Two factor analyses were then carried out. In the first a total of four factors were extracted, explaining 62 per cent of the variance. This ensured that the Boehm scores loaded on one of the Adjustment to School factors. After rotation by the Varimax procedure, the first three factors were identical with those identified when the scale was analysed alone. Factor 4, however, now showed substantial loadings for the Boehm scores. This result confirmed the identification of the factor as verbal ability.

In a second factor analysis five factors were extracted (explaining approximately 67 per cent of the variance). This had the effect of creating a factor on which the Boehm scores alone loaded more than 0·40. The Adjustment to School factors remain similar to those originally extracted, except that Item 7, 'reaction to strange adults', now loads with the items comprising the original Factor 4 and not Factor 2.

This separation out of the two instruments is most probably due to differences between the responses required, and indeed the respondents. Whereas the Adjustment to School Scale required teachers to rate children on a number of four-point scales, the Boehm test required five-year-old pupils to respond to a series of multiple-choice problems: it is, however, surprising that the separation is so complete; neither of the Boehm scores have loadings above 0·27 on any of the Adjustment to School factors.

Thus ability, as measured by the Boehm Test, is related to Adjustment to School, specifically to ratings of items concerned with verbal ability. Extraction of an extra factor, however, increases the explained variance from 62 per cent to 67 per cent and creates a separate factor for the Boehm scores.

Table 5: Comparison of factor scores for children with and without nursery experience—Pilot sample[1]

Factor No.	*Non-nursery* N	Mean	sd	*Nursery* N	Mean	sd	t	*Significance*[2]
1	47	23·23	6·98	78	20·86	6·60	1·88	$p < \cdot 05$
2	47	8·64	3·26	76	7·74	3·05	1·58	ns
3	51	12·10	2·18	83	11·87	2·21	0·58	ns

[1] Differences in numbers on the three factors are a result of missing data, i.e. teachers not rating children on particular items. A lower score indicates a superior rating.

[2] Significances stated are for a one-tailed test.

5. Comparison of children with and without experience in a nursery school or class

(a) *Pilot sample*

Factor scores were computed for all pupils, and the results for children with and without nursery experience compared using a t-test. As can be seen from Table 5, the value of t does not reach significance for any of the comparisons if the stringent two-tailed test of significance is applied. If, however, a one-tailed test is considered appropriate in the light of the general prediction that nursery children would be superior, the differences on Factor 1 reach significance at the 0·05 level.

(b) *Slough sample*

A similar set of comparisons carried out on the four factor scores obtained for the Slough sample reveals no differences between nursery and non-nursery groups (see Table 6).

Two possible explanations can be given for this contradictory result:

(i) Sixteen per cent of the Slough nursery children had experienced a special compensatory programme while in the nursery. Results for these children show them to receive less favourable ratings than either of the other groups on Factors 1, 3 and 4. These findings are contrary to what would be predicted. They may be a result of the way the children were selected to take part in the programme. Using the criteria devised by the National Children's Bureau (Joseph and Parfit, 1972), 80 per cent of these children are classified as being in special need, as compared with 50 per cent of a comparison group tested at the same time.

(ii) The Direct Entry group included 46 children from a school with a largely middle-class intake. If their ratings are looked at in comparison with the other non-nursery children, they are found to have significantly more favourable ratings on Factors 1, 3 and 4. If after excluding them the nursery/non-nursery comparisons are re-calculated, the nursery children are found to be significantly superior to the non-nursery children on Factor 4, verbal ability, and the difference on Factor 1 reaches significance if a less stringent criteria is used. This would suggest that a dif-

Table 6: Comparison of factor scores for children with and without nursery experience—Slough sample

Factor No.	*Non-nursery* N	Mean	sd	*Nursery* N	Mean	sd	*t*	*Significance*
1	192	19·83	5·99	147	19·40	6·07	$<1\cdot 0$	ns
2	199	5·34	2·01	150	5·19	1·93	$<1\cdot 0$	ns
3	191	7·71	1·85	142	7·70	1·89	$<1\cdot 0$	ns
4	196	4·96	1·94	151	4·73	1·74	$<1\cdot 0$	ns

Table 7: Results for pilot sample when factor scores are calculated on the basis of the Slough analysis

Factor No.	*Non-nursery* N	Mean	sd	*Nursery* N	Mean	sd	*t*	*Signifi-cance*
1	47	22·36	6·65	73	19·86	6·52	2·04	$p < {\cdot}05$
2	46	6·33	2·68	70	5·64	2·43	1·42	ns
3	50	8·28	1·85	76	7·78	1·82	1·51	ns
4	49	5·94	1·09	76	5·07	1·92	2·40	$p < {\cdot}05$

ference in social class composition is at the root of the more favourable scores of the non-nursery group.[1]

(c) *Comparison of the results for the two samples*

It was considered possible that the difference between the results for the two samples was attributable to the differently derived factor scores. Scores were therefore computed for the pilot sample on the same basis as those for the Slough sample. When comparisons between nursery and non-nursery children are carried out on these new scores, the results shown in Table 7 obtain. The differences between nursery and non-nursery children in the pilot sample are thus increased as the groups now differ significantly on Factors 1 and 4 ($p < {\cdot}05$, two-tailed).

When the mean factor scores for the pilot and Slough samples are compared, a tendency is found for the Slough teachers to rate their pupils more favourably than the pilot teachers. The differences in favour of the Slough group are statistically significant on Factors 2 and 4 and there is a trend in the same direction on both the other factors. Closer examination of the results shows that the difference is greatest between the two groups of non-nursery children (significant on Factors 1, 2 and 4).

As suggested when discussing the Slough data, this may be due in part to differences in social class composition. However, even when the known middle-class group is excluded from the Slough sample, the ratings remain more favourable than those for the pilot group.

The information available does not allow us to do more than speculate about other explanations. One possibility, is that the teachers in the two areas have different expectations of children starting school, and hence are rating them against different criteria. Another is that there are differences between the two areas in the ways in which children are prepared for, and introduced to, school. For example, the Slough schools would appear to cater for families belonging to more coherent and stable communities.

6. Summary and conclusions

(a) Since a positive relation has been found between school readiness and later educational performance, it was decided to develop an instrument to measure initial adjustment to school. This could then be used to explore the types of background and experience which make it probable that the child's initial experience of school will be a satisfactory one.

(b) The instrument developed consists of twenty four-point scales covering the areas of:

(i) personal emotional adjustment;
(ii) social adjustment;
(iii) attitudes and behaviour in response to the intellectual demands of the school.

This is intended to be completed by the teacher in the second half of the child's first term at school.

(c) Two samples of children were rated on the instrument and the results factor analysed. Overall, the scales fell into the predicted three clusters. On the second analysis, however, an extra factor concerned with verbal skill was extracted.

Intercorrelation of factor scores (Table 8) for part of the sample shows that these aspects of adjustment to school are not independent. Hence it might be worth investigating the use of an overall adjustment score.

---

[1] Analysis of the data for those children for whom social class information is available reveals a significant relation between social class and Factor 1.

Table 8: Inter-correlation of Adjustment to School Factor Scores

Factors	*N*	*Correlation*
F1, F2	152	0·43
F1, F3	142	–0·43
F1, F4	153	0·63
F2, F3	143	–0·13
F2, F4	159	0·43
F3, F4	144	–0·31

(d) Factors obtained from the two analyses were compared using a coefficient of congruence. The values obtained indicate a high degree of consistency.

(e) Scores on the Boehm Test of Basic Concepts were analysed with the Adjustment to School scales. If four factors only were extracted, the scores from this test loaded on Factor 4, verbal skill. Extraction of a fifth factor, however, resulted in a separate factor on which only the Boehm scores had high loadings, probably reflecting the differences in both the types of responses and the respondents.

(f) One type of early experience which is usually thought to play a positive part in preparing the child for school, is attendance at a nursery school or class. The results for both samples were analysed to test the prediction that when the total intake into a number of schools was rated on initial adjustment, the children who had nursery experience would be superior in all areas to those who had not. The results of the pilot study confirmed the prediction of more favourable ratings within the cognitive area, but not where the other two areas were concerned. No significant differences were found for the second sample, possibly due to the higher social class of the direct entry children.

(g) The children rated in the second sample have been followed up until their sixth term of infant schooling, and it is hoped to report the predictive validity of the Adjustment to School scale in the future.

(h) Clearly there are a number of further uses to which the scale can be put. For example, it could be used to compare:

(i) children from a variety of backgrounds; or
(ii) children whose parents have had special training;

and also to investigate the ways in which teachers' expectations relate to their rating of children's adjustment.

References

AUSTIN, J. J. and LAFFERTY, J. C. (Eds) (1968) *Ready or Not? The School Readiness Checklist Handbook*. Muskegon, Michigan: Research Concepts, Test Maker Inc.

BEREITER, C. (1972) 'An Academic Pre-school for Disadvantaged Children: Conclusions from Evaluation Studies'. In: STANLEY, J. C. (Ed.) *Pre-school Programs for the Disadvantaged. Five Experimental Approaches to Early Childhood Education*. Baltimore: Johns Hopkins University Press.

BOEHM, A. E. (1970) *Boehm Test of Basic Concepts*. New York: Psychological Corporation.

COHEN, L. and BAGSHAW, D. (1973) 'A comparison of the achievements of nursery school and non-nursery school children', *Durham Res. Rev.*, **VI**, 30, 735–41.

DOLL, E. A. (1965) *The Measurement of Social Competence: a Manual for the Vineland Social Maturity Scale*. Circle Pines, Minnesota: American Guidance Service.

DOUGLAS, J. W. B. and ROSS, J. M. (1964) 'The later educational progress and emotional adjustment of children who went to nursery schools or classes', *Educ. Res.*, **VII**, 1, 73–9.

GUNZBERG, H. C. (1969) *Progress Assessment Chart of Social Development* (Manual). Birmingham: SEFA Publications.

HARMAN, H. H. (1960) *Modern Factor Analysis*. Chicago: University of Chicago Press, pp. 256–9.

HARROLD, M. V. and TEMPLE, M. H. (1959–60) 'A study of children in the admission classes of four infant schools, making a comparison between those who have attended a nursery school and those admitted direct from home'. Unpublished thesis, Child Development Centre, Institute of Education, University of London.

HUSEN, T. (1962) (2nd edn.) *Pedagogisk Psykologi*. Stockholm: Svenska Bokforlaget.

ILG, F. L. and AMES, L. B. (1965) *School Readiness*. Behaviour tests used at the Gesell Institute. New York: Harper & Row.

JOHANSSON, B. A. (1965) *Criteria of School Readiness*. Stockholm Studies in Educational Psychology 9. Stockholm: Almqvist & Wiksell.

JOSEPH, A. and PARFIT, J. (1972) *Playgroups in an Area of Social Need*. Slough: NFER.

MCKENNELL, A. (1970) 'Attitude measurement: use of co-efficient alpha with cluster or factor analysis', *J. Soc.*, 4, 2.

STUKAT, K.-G. (1963) 'Effekter av lekskolevistelse pa barns utveckling. En evalueringsstudie av produkttyp'. Rapporter fran Pedagogiska institutionen. Goteborgs universitet. Quoted in: JOHANSSON (1965). *Criteria of School Readiness*. Stockholm Studies in Educational Psychology 9. Stockholm: Almqvist & Wiksell.

TAYLOR, P. H., EXON, C. and HOLLEY, B. (1972) *A Study of Nursery Education*. Schools Council Working Paper No. 41. London: Evans/Methuen.

WEIKART, D. (1972) 'Relationship of Curriculum, Teaching and Learning in Pre-school Education. In: STANLEY, J. C. (Ed.) *Pre-school Programs for the Disadvantaged. Five Experimental Approaches to Early Childhood Education*. Baltimore: Johns Hopkins University Press.

WIDLAKE, P. (1971) 'Does nursery education work?' *Child Education*, April.

The Prevalence of Behaviour Problems Among Children in Care

Bertram A. Akhurst, formerly Centre for the Study of Learning Difficulties, Goldsmiths' College

Summary

The Rutter Scales were used to assess the level of behaviour disorders reported by houseparents and teachers, in a group of 182 children undergoing long-term residential care in 62 children's homes run by local authorities and voluntary bodies. The percentage of children scoring above cut-off point for maladjustment was significantly higher than that for comparison groups drawn from the classes attended by the children in care and that found in previous studies of the general child population. The possible implications of diagnostic groupings, patterns of reporting and child care variables are discussed.

Introduction

Although surveys of behaviour problems can supply us with prevalence rates for the total population, the disorders are not distributed evenly and it is important, if we are to provide adequate and appropriate services, to have detailed knowledge of any groups particularly at risk. Previous studies have already shown the vulnerability, for instance, of educationally subnormal children (Chazan, 1964) and children with physical disorders (Rutter, Tizard and Whitmore, 1970) and there have been a number of studies suggesting a high prevalence of behaviour disorder among children in care.

The child care population is a very heterogeneous group both in terms of the circumstances causing admission to care and in the duration and kinds of care provided. Although very few of the children are orphans all have experienced varying degrees of separation from their families, friends and neighbourhood. They have also been shown to suffer a level of physical, social and educational handicap which is considerably higher than that of the general population (Mapstone, 1969). Most are admitted to residential care for a relatively short period of time (between 20% and 30% of the children in the Homes we visited had been in their current Home for less than a year) and those who do remain longer may be fostered or cared for in one of a wide variety of local authority or voluntary Homes ranging from small independent units containing six children to Cottage Communities for over two hundred.

This heterogeneity of population has led to considerable variations between previous estimates of the prevalence of maladjustment. Lewis (1954) for instance found a very high incidence of disturbance (75%) but was studying children in a reception centre where the effects of separation and the circumstances precipitating this could be expected to be most marked. Kellmer Pringle and Bossio (1965) however, who studied 142 children in a village community found that 33% of the children were reported as maladjusted by the teachers and 30% by the houseparents. Less than 40% were regarded as stable by each. More recently, Rees and Morley (1971) found that teachers reported three times the level of maladjustment among children in care compared with a control group drawn from the same schools.

Acknowledgements
I am grateful to the Department of Health and Social Security who support the project of which this study forms part. They are not responsible, however, for any of the results or views expressed. I also appreciate the helpful advice given by Professor J. Tizard and Mr W. Yule. Two Research Officers, Mr Robert Rider and Miss Barbara Schneider, carried out most of the field work. Miss Marilyn Farquhar assisted with the analysis of data and Miss Ruth Davis managed the secretarial work. The research staff are particularly grateful to the administrative, child care and teaching staff who responded so willingly to our inquiries.

The current study which forms part of a much wider inquiry into long-term residential care attempts to provide prevalence rates for behaviour disturbance based upon a widely selected sample of children undergoing long-term residential care in thirty-six local authority Homes run by London boroughs and twenty-six voluntary Homes situated in South-East England. Characteristic patterns of behaviour disorder are also described and the reports of teachers and houseparents compared.

Method

The Homes in which the children lived ranged from cottage communities to small independent units of six children. No short-stay, single sex or specialized units were included. The thirty-six local authority Homes were run by seventeen different London boroughs and were situated mainly in the Metropolitan Area and most of the twenty-six voluntary Homes, selected from seventeen voluntary organizations, were situated outside London in the Home Counties.

The 182 children studied individually were drawn from junior (9–10 y. 11 m.) and a senior (13–14 y. 11 m.) age group. They had been in their current Home for at least six months (and usually more than a year), were not suffering from a statutory handicap or undergoing psychiatric investigation or treatment and had no immediate prospect of restoration or transfer to an alternative form of care. In the Homes we visited, 13% of the children had been classed as handicapped, including 4% who were deemed maladjusted. Our study was concerned with those who had no formally recognized handicap.

The scales used were a later version of those employed by Rutter during the Isle of Wight Study (Rutter, Tizard and Whitmore, 1970). The scales are straightforward and quick to complete and afford a basis for calculating the prevalence of maladjustment and distinguishing broad diagnostic categories. Houseparents completed the A2 and teachers the B2. In order not to draw attention to the child care sample and also to provide a comparison group, teachers in those schools which did not form part of a children's Home also reported on a sample of children in the same class as the child in care. These were pre-selected on the basis of the initial letter of their surname and excluded children who were handicapped or in care.

Results

a. Prevalence

Table 1 shows the number and percentage of children of each sex and age group falling at or above the cut-off point for maladjustment on each scale in both the child care and comparison groups. As assessed by teachers, the percentage of children in care falling at or above the cut-off is consistently higher than that for the comparison sample. With the exception of the senior girls it is twice as high. Boys have a higher rate of reported disturbance than girls and that for the junior boys is higher than that for the seniors.

Table 1: Proportion of children falling above cut-off points for maladjustment (combined local authority and voluntary service)

A. Houseparent's A2 Scale		N ⩾ 13	% ⩾ 13
Juniors	Boys (59)	22	37·3
	Girls (47)	11	23·4
Seniors	Boys (43)	12	27·9
	Girls (33)	7	21·2

B. Teacher's B2 Scale		N ⩾ 9	% ⩾ 9
Juniors	Boys (55)	30	54·5
	Girls (42)	13	31·0
Seniors	Boys (40)	17	42·5
	Girls (30)	9	30·0

C. Comparison Sample: Teacher's B2 Scale		N ⩾ 9	% ⩾ 9
Juniors	Boys (127)	34	26·8
	Girls (118)	17	14·4
Seniors	Boys (107)	32	20·9
	Girls (81)	20	24·7

The results for the junior children in our sample are compared, in Table 2, with other studies of children of a broadly similar age range in the

Table 2: The proportion of junior children in the current sample falling above the cut-off for maladjustment on the teacher's B2 scale compared with previous studies of children in the general child population

	Boys			Girls		
Group	N	No. > 9	% > 9	N	No. > 9	% > 9
Current Sample (9–10 y. 11 m.) B2	55	30	54·5	42	13	31·0
Current Comp. Grp. (9–10 y. 11 m.) B2	127	34	26·8	118	17	14·4
Isle of Wight Study (9–11 y.) B2	1,743	175	10·0	1,683	85	5·6
London Borough (10–10y 11 m.) B2 (Rutter *et al*)	1,179	344	29·2	1,102	185	16·8

general population. It will be seen that although the percentage above the cut-off is over five times that found on the Isle of Wight for both boys and girls, a more recent epidemiological study of children in a London borough (Rutter, *et al.*, 1974) produced prevalence rates which resembled those of our own comparison group. Nevertheless, the rates for our child care sample are between two and three times as high.

A recent study of two large cottage communities (Yule and Raynes, 1972) produced almost identical mean levels of reported maladjustment for the combined age and sex groupings (Table 3).

Table 3: Mean level of reported maladjustment: current study and Yule and Raynes' (1972) study compared

	% ⩾ *Cut-off for Maladjustment*			
	(N=)	A2	(N=)	B2
Current Study	(182)	28·6	(167)	41·3
Group Cottage Study	(326)	27·3	(326)	38·0

These figures also compare reasonably well with those of Kellmer Pringle and Bossio (1965) using the Bristol Social Adjustment Guide. Applying Rutter's method of assessing 'true prevalence' of disorder[1] to those children reported on both scales we obtain an estimate of prevalence of 35·5%. In evaluating this estimate, however, note should be taken of the reservations made later in this paper.

b. *Diagnostic categories*

On the 'anti-social' and 'neurotic' subscales embodied in the Rutter Scales, there was a consistent trend for all groups of children (including the comparison groups), to have higher anti-social scores or equal scores on both subscales, rather than higher neurotic scores. The problems of children with equal scores, or 'mixed disorders' as Rutter calls them, were found during the Isle of Wight study to be more characteristic of anti-social than neurotic behaviour. This preponderance of anti-social disorders was noted among boys during the Isle of Wight study but, in contrast to our own results, neurotic rather than anti-social disorders were slightly more preponderant among the girls (cf. Table 4).

---

[1] During the Isle of Wight study psychiatric investigation confirmed as maladjusted:

50·4% of those children reported by parents alone;
40·8% of those children reported by teachers alone;
73·7% of those children reported on both scales.

The resultant figure was subjected to a correction factor $\left(\frac{5}{4}\right)$ to compensate for known false negatives.

Table 4: Diagnostic categories (derived from sub-scales) of children falling at or above the cut-off for maladjustment on either the A2 or B2 scale

	A2			B2		
	N	*A/S and Mixed*	*Neurotic*	*N*	*A/S and Mixed*	*Neurotic*
Boys	(34)	25	9	(47)	39	8
Girls	(18)	14	4	(22)	14	8

c. *Symptoms*

To permit comparisons with other results, the prevalence of individual symptoms was calculated independently of their rate of severity. There were relatively few significant differences on items common to the A2 and B2 scales. Among the boys, teachers reported more restlessness, and inability to settle and the houseparents more fearfulness. Houseparents also reported more bullying in the seniors. All the significant differences among the girls were in the direction of a higher prevalence on the houseparent scale. They reported more nail-biting among the juniors, stealing and tics in the seniors and bullying in both age groups. Differences were significant at the 0·05 level only.

In their report of the Isle of Wight study, Rutter, Tizard and Whitmore provide tables showing the frequency with which individual symptoms are reported in the general child population studied and in a smaller sub-group of children diagnosed as disturbed as a result of psychiatric examination. These tables enabled our junior child care groups to be compared with both sets of data to discover which items had prevalence rates which resembled the psychiatric group rather than the general population, i.e. those symptoms where the prevalence rate was significantly higher than that for the general population ($p \leqslant 0{\cdot}01$) but not significantly different from the psychiatric group on both houseparent (A2) and teacher (B2) scales. Taking all children in the junior groups (and not just those who fall above the cut-off for maladjustment), the girls resembled the psychiatric group on 'fidgetiness' and all but one of the 'anti-social' items whereas the boys resembled the psychiatric group in 'destructiveness' only.

The child care group differed significantly ($p \leqslant 0{\cdot}01$) from the comparison group in the prevalence of relatively few symptoms on the teacher's scale. The junior boys were reported significantly more frequently on all the anti-social items and were more 'resentful of correction' and 'irritable'. The latter two symptoms were also significantly more common among the junior girls, as was 'restlessness' and 'disobedience'. There were no differences at this level of significance ($\leqslant 0{\cdot}01$) in the senior groups. The only item which was reported consistently and significantly less frequently for the child care groups was 'absence from school for trivial reasons'.

Anti-social behaviour is particularly characteristic of the junior boys for whereas all the items which discriminate anti-social behaviour on the houseparent's scale and all but one on the teacher's scale occur in at least one-third of the children in this group, only one such item reaches this level of prevalence on the houseparent's scale and none on the teacher's scale in the group of senior boys. The girls showed far fewer anti-social symptoms at this level of prevalence. 'Disobedience' was the only such item reported; for the juniors on both scales and, for the seniors on the teacher's scale only. Houseparents, however, reported an increased prevalence of anti-social items among the senior girls.

Of the neurotic items, 'worrying' was reported in more than one-third of the children in each age and sex group but it was also prevalent in all but the senior boys, in the comparison group. In all the child care groups, 'fearfulness' reaches this level and, among the senior girls on the teacher's scale 'miserable' and 'unhappy' as well.

d. *Agreement between scales*

Table 5: Agreement between houseparent (A2) and teacher (B2) scores

		Houseparent's (A2) Scale	
		Not Maladjusted (score <13)	Maladjusted (score $\geqslant 13$)
Teacher's (B2) Scale	Not Maladjusted (score <9)	76	22
	Maladjusted (score $\geqslant 9$)	44	25

The extent to which houseparents and teachers agree in diagnosing maladjustment in terms of the scores on their respective Rutter scales is no greater than could have been expected by chance (χ^2 [Yates] $= 3{\cdot}15$) and, as in previous studies, the A2 and B2 scales tend to select different groups of disturbed children (Table 6). Nevertheless, the level of agreement between the ratings of teachers and houseparents was four times that found during the Isle of Wight study in which natural parents completed the A2 scales. Teachers also reported a significantly higher proportion of disturbed children than did houseparents (χ^2 [Yates] $= 5{\cdot}82$ $p \leqslant 0{\cdot}02$).

c. *Child care variables*

Although only the scores of junior boys on the A2 scale showed a significant correlation with age admitted into care ($r = 0{\cdot}32$, $p \leqslant 0{\cdot}05$) and total time in care ($r = 0{\cdot}35$, $p \leqslant 0{\cdot}01$) a significantly higher proportion of children (68%) admitted below the age of five were scored at or above the cut-off maladjustment on either the houseparent's (A2) or teacher's (B2) scale compared with those who were admitted at a later age (46%) ($\chi^2 = 7{\cdot}381$, $p \leqslant 0{\cdot}01$). Taken separately the difference was only significant for the junior group although the trend was the same for the seniors.

Both the prevalence of children scoring above the level of maladjustment and the total scores for the various age and sex groupings suggest that, as a group, the junior boys in local authority Homes, were more disturbed than those in the voluntary Homes. The relation of child care variables to maladjustment levels is the subject of continuing inquiries.

Discussion

Our findings indicate that children undergoing long-term residential care have a significantly higher prevalence of severe behaviour disorder than has been reported in the general population and our results conform closely to those of a number of previous studies of children in care.

If we accept Rutter's formulation of the A2 and B2 scale results a valid method of estimating prevalence of disorder in our study, the result is still likely to underestimate the true prevalence rate. In selecting our sample we specifically avoided the 13% of the child care population of the Homes visited who had been formally ascertained as handicapped, and this included 4% who had been deemed maladjusted. The gross figure also obscures the relatively greater degree of disturbance found among junior boys in the more numerous local authority compared with voluntary Homes.

The validity of using Rutter's estimate of prevalance in this study must also be questioned. Our sample had a wider age range than that assessed during the Isle of Wight Study (Rutter, Tizard and Whitmore, 1970) and, unlike that study, we did not have the benefit of an independent psychiatric validation of our findings. Houseparents' reports could, for instance, relate to psychiatric opinion in a different way from those of true parents. Any estimate of prevalence must, therefore, be interpreted with caution.

The discrepancy between the levels of reporting in our child care groups and the comparison groups drawn from the classes they attend could, it is suggested, underestimate the higher prevalence of disorder among the children in care. If the children in care are more likely to present behaviour difficulties (and perhaps are more retarded educationally as previous studies suggest [Kellmer Pringle and Bossio, 1965]) then there may be a greater chance that other children in their class also present such difficulties. Thus, the comparison groups cannot be equated with the general child population.

The high prevalence of anti-social behaviour warrants further discussion and inquiry. Such problems were prevalent in the teachers' reports on children in the comparison groups as well as the children in care and it is also possible that houseparents would report anti-social behaviour more readily than would a child's natural parents. The seriousness of anti-social behaviour has been emphasized by studies suggesting that it is particularly resistent to current methods of psychotherapy (Rachman, 1971) and that it may be transmitted culturally to successive generations (Wardle, 1961).

Our results do not, at this stage of the analysis, throw any light on the reasons for the high prevalence of disorder among children in care, its relation to early admission and the possible decline in anti-social behaviour with age. A preliminary examination of the case histories, however, reveals common patterns of experience which, by any standards, would be regarded as disturbing even without recourse to the effects of separation itself. It would also be consistent with current studies of child development to expect that their effects would be more severe and persistent when they occurred in the developmental periods of greatest psychological and physical maternal dependency. The care provided in most of the Homes we visited was far removed from the stereotype of past institutions and we could well believe that the prolonged and consistent experience of such supportive care could account for the remission of some of the symptoms such as were noted in the older boys. However, as yet, we have no evidence for or against this view. It is known that many childhood problems tend to remit with age (Shepherd, *et al.*, 1971) and although our senior boys did show fewer anti-social problems than the juniors, there was a smaller proportion of early admissions to care in the senior group (although the difference was not statistically significant).

References

CHAZAN, M. (1964). 'The incidence and nature of maladjustment among children in schools for the educationally subnormal', *Brit. J. Educ. Psychol.*, **34**, 292–304.

KELLMER PRINGLE, M. L. and BOSSIO, V. (1965). 'Intellectual, Emotional and Social Development of Deprived Children'. In: KELLMER PRINGLE, M. L. (Ed.) *Deprivation and Education*. London: Longman, Ch. 2.

LEWIS, H. (1954). *Deprived Children*. London: Oxford University Press.

MAPSTONE, E. (1969). *Children in Care*. Concern, No. 3. London: National Children's Bureau.

RACHMAN, S. (1971). *The Effects of Psychotherapy*. International Series of Monographs in Experimental Psychology, Vol. 15. Oxford: Pergamon.

REES, H. M. N. and MORLEY, R. (1971). 'An epidemiological survey of psychiatric disorders of children in care'. Not yet published.

RUTTER, M., COX, A., TUPLING, C., BERGER, M. and YULE, W. (1974). 'Attainment and adjustment in two geographical areas. I. The prevalence of psychiatric disorder', *Brit. J. Psychiat.*, December.

RUTTER, M., TIZARD, J. and WHITMORE, K. (1970). *Education, Health and Behaviour*. London: Longman.

SHEPHERD, M., OPPENHEIM, B., and MITCHELL, S. (1971). *Childhood Behaviour and Mental Health*. London: University of London Press.

WARDLE, C. J. (1961). 'Two generations of broken homes in the genesis of conduct and behaviour disorders in childhood', *Brit. Med. J.*, **II**, 2, 349–54.

YULE, W. and RAYNES, N. (1972). 'Behavioural characteristics of children in residential care in relation to indices of separation', *J. Child. Psychol. Psychiat.*, **13**, 249–58.

British Journal of Educational Psychology

Prof. N. J. ENTWISTLE, Department of Educational Research, University of Lancaster

Vol. 44 Part 3 November 1974

J. D. NISBET: Editing the Journal.
M. H. NEVILLE and A. K. PUGH: Context in Reading and Listening: a Comparison of Children's Errors in Cloze Tests.
G. W. HERBERT: Teachers Ratings of Classroom Behaviour: Factorial Structure.
A. G. SMITHERS and S. DANN: Success and Failure among Engineers, Physical Scientists and Linguists at a Technological University.
H. PHILP and M. KELLY: Product and Process in Cognitive Development: Some Comparative Data on the Performance of School Age Children in Different Cultures.
S. C. SOMERVILLE: The Pendulum Problem: Patterns of Performance Defining Developmental Stages.
R. D. DUTCH and J. MCCALL: Effects of Introducing a Transition Department in a Secondary School.
H. FRANCIS: Social Background, Speech and Learning to Read.

RESEARCH NOTES:
R. J. RIDING and J. M. SHORE: A Comparison of Two Methods of Improving Prose Comprehension in Educationally Sub-Normal Children.
A. G. SMITHERS and D. CHILD: Convergers and Divergers: Different Forms of Neuroticism?
P. M. MILLER and R. R. DALE: A Comparison of the Degree Results of University Students from Co-educational and Single-sex Schools.
C. EISER: Recognition and Inference in the Co-ordination of Perspectives.
D. BOYDELL: Teacher-Pupil Contact in Junior Classrooms.

BOOK REVIEWS.

PUBLICATIONS RECEIVED.

INDEX TO VOL. 44.

NOTES TO CONTRIBUTORS.

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Junior School Pupils' Rejection of School Library Books

G. Fenwick, Ethel Wormald College of Education

Summary

One hundred and ninety one children, aged ten and eleven, kept records for a period of six weeks of the fiction books which they selected from their school libraries. These records indicated whether books were completely or partly read (i.e. rejected) before they were returned to the libraries.

When factors related to the children and their environment were examined, the results showed social class and the provision of school library books to be linked significantly with this type of book rejection. An examination of a sample of the books selected revealed a significant relationship between rejection and the number of words and proportion of illustration which a book contained.

Introduction

It is common knowledge that children are by no means certain to read completely all of the books which they select from school and class libraries and it would be naive to think that most teachers are not aware that this type of rejection exists.

How teachers react to this type of rejection will vary from individual to individual but except when rejection of fiction books goes unnoticed there appear to be three likely approaches to the problem. Firstly, rejection might be regarded as a normal occurrence to be actively encouraged or, at least, not to be discouraged. Secondly, it might be regarded as a common occurrence which should, nevertheless, be discouraged. Thirdly, the aim might be to prevent rejection from taking place.

Attempts to encourage, discourage or tolerate children's rejection of fiction books are at the present time highly individualistic and lacking in guidance because there has been no investigation of likely causal factors. The nearest approaches have been those of Bamberger (1922) and Carsley (1957), both of which examined the make up and appearance of books to determine how children selected them from the library shelves. No research can be traced which investigates reasons for children rejecting fiction books once they have been taken from the library.

To claim that a book is either accepted or rejected might well be creating too coarse a division because the practice of obtaining the general gist of a story without actually reading every page is not allowed for. It can be argued that this practice, commonly termed 'skipping', is more akin to acceptance than to rejection in that a general understanding of the story is probably obtained. For the purpose of this study,[1] is was regarded as being part of the process of acceptance.

The hypotheses

A number of hypotheses were tested in an attempt to discover the pattern and nature of children's book rejection. They related to: 1. the children and the conditions under which they read; 2. factors inherent in the books which they selected.

The hypotheses were as follows:

1. There will be no marked difference in the percentage of books rejected by boys and girls respectively.

2. The percentage of books rejected by individual children will vary with social class.

3. There will be an association between a child's high percentage of rejection and certain home circumstances, viz: (i) few children's books in the home; (ii) children's non-membership of public libraries.

4. The personality factors of 'Instability',

---

[1] This paper reports work of an investigation carried out for the MEd degree at the University of Manchester and detailed in an unpublished thesis, 'Junior school pupils' rejection of school library books'.

'Expedience' and 'Extraversion' will be associated with a child's high percentage of rejection.

5. A child's high percentage of rejection will be associated with a lower than average reading ability.

6. Children's percentage of rejection will be highest in schools where the provision of school library books is small.

7. Children's percentage of rejection will be highest in classes where the class teachers score highly for the personality factors of 'Naturalism', 'Radicalism' and 'Tendermindedness'.

8. Rejected books will contain, on average, more words than accepted books.

9. Rejected books will contain, on average, a smaller percentage of page area given to illustration than will accepted books.

10. The reading ease and human interest of rejected books will be less, on average, than that for accepted books.

The sample

Six primary schools, all of them possessing libraries containing at least 1,000 books, were used. The schools were divided into three fairly crude socio-economic categories, namely urban, estate and suburban. There were two schools in each category.

The work was carried out with unstreamed classes of fourth-year junior children aged ten and eleven years. A number of researchers, in particular Terman and Lima (1931) and Leng (1968), have found that interest in reading tends to be at its most intense at about this age. It was thought that the reading ability and powers of discrimination of these children in relation to the selection and use of literature would be sufficiently well developed to provide results within a relatively short period of time.

Procedure

Each child filled in a record sheet for every fiction book borrowed from the school library during a period of six weeks. From the record sheets it was possible to discover how many books each child accepted and rejected. From this information a child's percentage of rejection was calculated by expressing the number of books rejected as a proportion of the number of books borrowed.

Measurement techniques

In addition to the record sheet, the following tests, scales and questionnaires were used.

To determine children's membership of public libraries and the number of books which they possessed at home, use was made of a home information sheet.

The Hall–Jones social grading of occupation scale was used to determine the social class of individual families. This scale is used fairly frequently in educational research. Its main advantage, when compared to the Registrar-General's scale is that it consists of seven different grades of social class as opposed to the latter's five.

Each child completed Cattell and Porter's children's personality questionnaire. Commonly known as the CPQ, this questionnaire gives a general assessment of fourteen distinct dimensions of personality for children aged between eight and twelve years.

Schonell's graded reading vocabulary test R1, a test frequently used in schools, was used to assess children's reading ages.

The teacher in charge of each class completed R. A. C. Oliver's Survey of Opinions about education. Using a factorial scoring key, scores were obtained for the personality dimensions of 'Naturalism', 'Radicalism' and 'Tenderminded-ness'.

The reading ease and human interest of a random sample of the books selected by the children were assessed by using the Flesch scales. The method of assessing reading ease takes into account two factors, these being the number of syllables in a word, and the number of words in a sentence. Human interest takes into account 'personal words', relating to people, and 'personal sentences', which in most cases consist of conversations.

Lastly, a library provision index was created because it was considered that an assessment of school libraries should take into account not only the number of books but also the number of children using them. In order to standardize assessment, the following formula was used:

$$\text{Library provision index} = \frac{\text{Number of books in school library}-1{,}000}{\text{Number of children who use school library.}}$$

One thousand books was taken, fairly arbitrarily, as being the absolute minimum number of books required by a school library to allow

Table 1: Number of children in various categories of rejection

	Absolute rejectors (100% rejection)	*High rejectors (99% to 51% rejection)*	*Mid-point rejectors (50% rejection)*	*Low rejectors (49% to 1% rejection)*	*Acceptors (0% rejection)*	*N*
Boys	10	4	10	25	34	83
Girls	10	16	8	34	40	108
	20	20	18	59	74	191

Chi^2 = 6·138; df ·4 Not significant

children a reasonably free choice of selection. This figure was rather low, especially in the light of Roberts' (1969) claim, and supporting evidence in the Plowden Report (Cace, 1967) that a primary school with an entry of one class per year requires at least two thousand library books in order to cover the children's wide range of interests.

There were wide fluctuations of library stock between schools. Several schools possessed over three thousand books while others had to be discarded because their library stock was so tiny. The age of a school and its environment appeared to determine, to some extent, the size of the library. It takes time to build up library stocks and this tends to happen more quickly in suburban, as opposed to urban environments. Taking into account these variables and the many other demands that the research was making upon the schools involved, the figure of one thousand was somewhat reluctantly accepted as being the minimum number of books with which a primary school library could expect to function efficiently.

Results

The 108 girls who completed record sheets selected 486 books, appreciably more than the 296 books selected by the 83 boys. The difference bears out the findings of Terman and Lima (1931) Jenkinson (1940) and Leng (1968), amongst many others, who have found that from an early age girls tend to read a greater quantity of fiction than do boys. Despite this, the percentage of rejection of boys and girls was not significantly different. Nor, as Table 1 shows, were there any significant differences when rejectors were placed in a number of categories. The patterns of rejection of boys and girls seem to be fairly similar.

The actual number of books selected by individual children during the investigation appeared to have little effect upon rates of rejection. Not surprisingly, children who did not reject at all tended to read fewer books than children who did reject, but not markedly so. While the number of books selected ranged from one to sixteen, the average for boys was four and that for girls five. As Table 2 indicates, the percentage of rejection was not greatly different for both boys and girls who selected either more or less than the average. This might well be due, to some extent, to the fact that the investigation lasted for only six weeks. One would suspect that a lengthier investigation might reveal greater variation in the relationship between quantity of selection and the percentage of rejection.

Table 2: Mean percentage of rejection related to number of books selected

	Number of books selected		
	Below mean	Mean	Above mean
Boys	31·0 (n = 45)	27·0 (n = 12)	34·5 (n = 26)
Girls	31·4 (n = 63)	29·3 (n = 15)	30·6 (n = 30)

Children who possessed few books at home or who were not members of public libraries did not reject a significantly higher proportion of books than children who enjoyed more favourable facilities. When these two factors were combined, however, the results approached significance at the ·05 level, giving some indication that family literacy might affect rejection. Children from homes where family literacy (as measured by public library membership and books in the home) was high tended to reject more frequently than did children from homes where family literacy was low.

There was no significant association between children's high rejection and the personality factors of Instability, Expedience and Extraversion save that the small number of children who rejected all the books which they selected had, on average, lower scores for all three traits than the rest of the children in the sample.

Lower than average readers (i.e. those with a reading age of less than 9½ years) did not reject significantly more books than average and above average readers, although it should be remembered that there were very few extremely slow readers in the sample. Class teachers who scored highly for the personality traits of 'Naturalism', 'Radicalism' and 'Tendermindedness' did not appear to influence their pupils to reject a higher proportion of books than pupils in classes where the teachers' scores were low.

Approximately 250 of the over 500 different books read by the children were analysed for reading ease and human interest. Analysis of variance showed no difference between books which were consistently rejected by their readers, books which were sometimes accepted and sometimes rejected, and books which were consistently accepted. This result must be treated with caution as users of the Flesch scales for reading ease and human interest are recommended to sample a large number of passages from each book and lack of time prevented such a large scale analysis. In most cases, no more than six samples per book were taken.

Four of the hypotheses produced statistically significant results. A significant difference at the ·05 level was found between rejection percentages when the children were divided by social class. In order to determine if there was any pattern to this difference the proportion of high rejectors (i.e. children who rejected 50% or more of books selected from school libraries) in each social class was compared and tested for trend. While half of the children in Social Classes 1 and 2 combined (professional and managerial) were high rejectors, the proportion declined steadily until approximately one eighth of the children in Social Class 7 (unskilled) were in that category. The trend, as shown in Table 3, was significant at the ·01 level.

When the six schools were arranged in order of library provision and the proportion of high rejectors in each school was tested for trend, the results for boys and girls were significant at the ·05 level (see Tables 4 and 5). As the results indicate, there were unlikely to be as many high rejectors in a school where the number of books available for selection was small in proportion to the number of children as there would be in a school where the number of books available was large. This evidence suggests the converse of hypothesis 6. A child's percentage of rejection was likely to be highest not in schools where the provision of library books was small, but in schools where it was great.

Table 3: High and low rejectors by social class, boys and girls combined

	Social Gradings						
	1 & 2 (combined)	3	4	5	6	7	N
High rejectors (Reject 50% or more of selected books)	9	9	12	17	7	4	58
Low rejectors (Reject 49% or less of selected books)	9	18	24	28	21	27	127
Total	18	27	36	45	28	31	185

Testing for trend

Source of variation	*df*	*chi*^{2}	*Level of significance*
Due to linear regression	1	6·710	$p < ·01$
Departure from regression line (by subtraction)	4	2·577	NS
Overall value	5	9·287	NS

Table 4: Number of high and low rejecting boys in schools arranged in order of library provision index

	Schools in order of library provision index (lowest first)						
	F	E	B	D	A	C	N
High rejectors	1	0	5	3	6	9	24
Low rejectors	6	12	7	11	11	12	59
Total	7	12	12	14	17	21	83

Testing for trend

Source of variation	*df*	*chi*^{2}	*Level of significance*
Due to linear regression	1	5·1500	$p < \cdot 05$
Departure from regression line (by subtraction)	4	4·1045	NS
Overall value	5	9·2545	NS

Table 5: Number of high and low rejecting girls in schools arranged in order of library provision index

	Schools in order of library provision index (lowest first)						
	F	E	B	D	A	C	N
High rejectors	3	2	6	11	4	8	34
Low rejectors	14	13	12	11	9	15	74
Total	17	15	18	22	13	23	108

Testing for trend	*df*	*chi*^{2}	*Level of significance*
Source of variation	1	4·066	$p < \cdot 05$
Departure from regression line	4	3·386	NS
Overall value	5	7·452	NS

When the number of words contained in books was analysed there was a significant difference at the ·01 level (Table 6). In all six schools the means for consistently rejected books were higher than for consistently accepted books. Books which were consistently accepted tended to be fairly short and were unlikely to contain more than 40,000 words. The proportion of illustration contained in books was also analysed and the result was significant at the ·01 level. Books which were consistently accepted tended to have a higher percentage of their content given to illustration.

Table 6: Mean number of words for three categories of books, school by school

Schools	*Rejected books*	*Rejected/ Accepted books*	*Accepted books*	*School means*
A	66,875	18,667	38,047	44,730
B	71,842	55,800	39,720	54,177
C	39,975	44,205	23,530	32,175
D	52,867	59,710	31,547	42,890
E	53,442	25,700	31,193	39,543
F	68,350	62,241	11,420	50,670
Overall Means	58,937	42,170	33,523	

Analysis of variance between three categories of books, $P < \cdot 01$

Discussion

The results of this investigation relate mainly to the age group of the children who were studied, that is, ten- and eleven-year-olds. The fairly gradual nature of the changes in children's reading tastes and habits suggests that they are also fairly relevant to children between the ages of eight and twelve. Although the relative smallness of the sample indicates that these results must be treated with caution they suggest, at the very least, areas where large scale investigations might be useful.

The investigation establishes, to some extent, that this type of book rejection occurs quite frequently (Table 1). Furthermore, there are some grounds for assuming that it is a useful factor in children's reading habits.

Poor readers do not reject a significantly higher percentage of books than do good readers, so rejection does not appear to reflect reading ability. But it does seem related to social class and the large amount of evidence reporting the disadvantages encountered by working class children during their formal education suggests that their tendency in this investigation, to reject a fairly small proportion of fiction books, is not to their advantage. This likelihood is reinforced by the fact that the two urban schools, F and E, where the pupils were predominantly working class, had the smallest library provision. Children who rejected highly tended to be those who enjoyed better school library facilities and who came from middle class backgrounds. The near significant result for family literacy tends to support this finding. High rejectors were more likely to have a fairly large stock of books at home and were more likely to be members of public libraries.

The apparent links between social class and library provision as factors of book rejection seem to indicate that positive discrimination of the type recommended by Plowden (Cace, 1967) might be needed in relation to the use and provision of children's literature in schools in urban areas.

Rejection might be seen as a positive factor of children's reading which enables them to be discriminating in their choice of fiction. The ability to discriminate in this way seems linked to the number of fiction books available. In a country where 3,000 children's books are published annually such discrimination is probably useful and perhaps even necessary, provided that this mass of literature is reflected in the stocks of school libraries. This positive value of rejection might be questioned in the case of the 20 children who rejected all of the books which they selected from their school libraries. While extreme caution must be exercised when examining such small numbers there are grounds for suspecting that children who are consistently high rejectors might be less stable, less dependable, and more introverted than other children. At the very least, there must be cause for concern on educational grounds when children reject virtually all available fiction out of hand.

When factors relating directly to the books selected are considered, there seems little doubt that children fight shy of lengthy books. Yet an examination of the school libraries showed that they contained a considerable number of books of 65,000 words or more. Many of these books were children's classics such as *Huckleberry Finn*, *Hans Andersen's Fairy Tales*, *Swallows and Amazons* and the *Hobbit*. Carsley (1957) pointed out that teachers depend quite heavily on children's classics when they select books for school libraries. Keeping up to date with the mass of fiction which is published annually is by no means easy and it seems likely that many school libraries contain fairly large numbers of books which have been selected not so much because children *do* read them as because teachers think that they *should* read them.

When the choice of fiction is so great and library allowances generally small, some attention should be given to the length of books which are selected for school libraries. Attention might also be paid to the second significant factor related to the actual books, namely, illustration. Popular books tend to be those having 10 % to 20% of their content taken up with illustrations. A book with less than 5% illustration in its content is unlikely to be very popular.

Generally, the significant evidence relating to books appears to be stronger than that relating to social class and library provisions. This might be seen to support Whitehead's (1956) contention that the attitudes of children towards books are influenced not so much by factors relating to the children themselves and the conditions under which they read as they are by factors inherent in the books.

In this investigation little account was taken of classroom and library organization. None of the schools operated a fully integrated day so the question of whether children working within such a system tend to have more time for individual reading remained unanswered. Time does, however, seem to be an important factor. Table 7 indicates that the majority of books which were completed were those which were read within a few days. This might well be one of the reasons why short books appeared to be so popular. It might also indicate that personal reading time should be allocated fairly generously in schools, particularly in those where children are not encouraged to take school library books home.

Table 7: Amount of time taken to complete books which were consistently accepted

	Time taken to complete book		
	A few days	More than a week	More than a fortnight
Number of books	102	23	8

Although school library organization was not subjected to close scrutiny, it was noticed that the school where library rules were least flexible was the one with the highest overall rates of rejection. Insistence that school library books must be returned from home weekly, on one particular day, might well produce a type of book rejection which is negative and of little use.

In the introduction to this report four possible approaches to rejection were mentioned. Rejection can be encouraged, tolerated, discouraged or prevented. This investigation appears to show that children's book rejection should be, at the very least, tolerated. One might go further and say that in fostering critical attitudes towards literature teachers might well encourage children to reject books which they do not particularly like. There are two exceptions to this rule.

Firstly, children who reject all, or almost all, of the books which they select need to be helped and encouraged to adopt a more positive approach. The results of this investigation suggest tentatively that some of these children might be experiencing personality problems. They might also be experiencing reading difficulties, for although there is no evidence that poor readers reject more books than average or good readers it should be remembered that the children in this investigation were fourth-year juniors and by the time that stage has been reached many very slow readers have been dispersed to special schools or classes for slow learners.

Secondly, a high overall rate of rejection within a class might sometimes be an indication of methods of library organization which are too rigid.

Whatever the individual teacher's attitude is towards rejection, its measurement is a useful form of feedback. The fairly simple form of measurement used in this investigation was generally thought to be useful by the teachers involved and proved to be popular with pupils.

References

BAMBERGER, F. E. (1922) *The Effect of Physical Make-up of Books Upon Children's Selection*. New York: Johns Hopkins.

CARSLEY, J. D. (1957) 'The interests of children (aged 10-11) in books', *Brit. Educ. Psychol.*, **27**, 13–23.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1967) *Children and their Primary Schools*. London: HMSO, Vol. I, p. 214.

JENKINSON, A. J. (1940) *What do Boys and Girls Read?* London: Methuen.

LENG, I. J. (1968) *Children in the Library*. Cardiff: University of Wales Press.

ROBERTS, G. (1969) *Reading in Primary Schools*. London: Routledge & Kegan Paul.

TERMAN, L. M. and LIMA, M. (1931) *Children's Reading*. New York: Appleton.

WHITEHEAD, F. S. (1956) 'The attitude of grammar school pupils towards some works of prose fiction commonly read in schools', *Brit. J. Educ. Psychol.*, **24**, 104–10.

Education 3-13

A journal for Primary Education

In the April issue

Editorial: Claims. *Colin Richards*

The beginning of a journey: a school opens. *Arthur Razzell*

Skill-centred teaching: an alternative to integration. *David Oliver*

Reading: seven to eleven. *John Merritt*

Spelling: further thoughts. *Margaret Peters*

You're fine. How am I?: status and the primary teacher. *Derek Sharples*

It may be ordered direct from the Publishers

In Britain

The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow G64 2PW

Annual subscription £1.50 including postage

In the U.S.A. and Canada

APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

School Inequality and the Welfare State

JOHN D. OWEN

This provocative analysis of the structure of educational inequality in America today is a major contribution to current discussion of school financing reforms.

While showing that sources of inequality are deeply imbedded in our present system, the book presents a far-from-hopeless view of the future. The author gives evidence of strong pressures for change which have developed in the past decade, and his appraisal of the long-term prospects for a reduction in schooling inequalities is a positive one.

John Owen points out that until recently the traditional decentralized system worked well, at least for the nonpoor, but suggests that the coming of the welfare state has radically altered the political economy of education. His specific proposals for democratic reform will be of interest to parent and taxpayer, as well as to educator and economist.

In documenting his case, Owen evaluates inequality within the school systems of a number of cities. He presents the first systematic national study of intercity and interstate differences in the ability of teachers and shows how our traditional system of college scholarships, loans, and subsidized tuition has often operated to assist the middle class to increase its educational advantage over the poor.

240 pages £5.50

Johns Hopkins University Press

2-4 Brook Street, London W1

Book Reviews

The English Sixth Form

by P. H. Taylor, W. A. Reid and B. J. Holley
Routledge & Kegan Paul, 1974, 186 pp.,
Hardback £3·95

'Schools commit themselves to sixth-form courses in the belief that these serve a purpose, a desired end, for a selected group of young men and women, in developing in them abilities and aptitudes which are valued both by them and by society, and are part of a worthy educational tradition. It is the empirical study of these beliefs and their implications in an institutional context and in relation to other organizations which is the substance of this book. It cannot claim to have researched all aspects of the sixth-form curriculum. But it can claim to have attempted to make explicit characteristics of the sixth-form curriculum which previously were unnoticed or neglected.'

It is rare for a book to sum itself up so neatly or so concisely as in the preceding paragraph (taken from the Introductory chapter). The English sixth-form and its curriculum are distinctly unusual; a fact most readily appreciated by those who have been involved with the secondary school systems of other developed countries. The sixth-form when measured by its own criteria is successful, remains comparatively unaffected by the pressures of society at large, is characterised by strong conservatism both among teachers and students, and yet, seemingly, always teeters on the point of revolutionary change.

The authors (from the Teaching Research Unit of the University of Birmingham) probe the different aspects of the sixth-form to considerable depth, basing their findings on a substantial mass of data collected from students, teachers, school heads, university selectors, etc. This is sandwiched between an excellent short history of the sixth-form, and a somewhat speculative description of sixth-form colleges and certain other contemporary developments. Most interesting perhaps are the data on teacher attitudes which begin to explain the remarkable resilience that the traditional sixth-form has exhibited in recent years.

The authors conclude, however, that changes *will* be introduced eventually. 'Sixth-forms as we have known them will stay with us for some time yet. Any policy decisions will take time to implement, but as from the late 1970's there is likely to be a progressive merging of the traditional sixth-form into a system of general provision for 16+ education. This will involve not so much a loss of identity for the sixth-form as its contribution to a new and necessary evolution of education for the new adults of a new era' (p. 154).

This is a fascinating book packed with interesting facts, seductive hypotheses, and explanations for most of the obscure ritual that has come to be associated with sixth-form. With a neat combination of large scale and small scale approaches, it succeeds in illuminating an important part of our educational system.

BRUCE CHOPPIN

Present and Future in Higher Education

edited by R. E. Bell and A. J. Youngson
Tavistock Publications Ltd., and Harper & Row Publishers Inc. 1973, 192 pp, Hardback £3·40

The Government's acceptance of the Robbins' principle that there should be places in higher education for all students who were qualified and wanted them heralded a long period of expansion and development in post-secondary education in Britain. Seven new universities have been created, the ten Colleges of Advanced Technology have been upgraded, thirty polytechnics have been formed, and the whole system of teacher training has been thrown into the melting pot. What was once a small, stable network of universities and colleges of various kinds, largely out of the public eye, has been pitchforked into national prominence and loaded with all kinds of expectations which have not been easy to meet.

During December 1970 the Faculty of Social Sciences in the University of Edinburgh held a Seminar to assess the situation as it appeared then and to speculate on the future, and out of this has grown *Present and Future in Higher Education.* There are papers which attempted to set British post-secondary education in its historical (Kearney, Bell) and international context (Grant, Watson), which look at aspects of the current scheme of things—for example, the new universities (Cottrell) and 'the other side of the binary system' (Lowe), which set out some hopes for the future (Wood, Thompson, Young, Wolfe), and which analyse some features of the planning and control of higher education (Barnes, Vaizey).

Many of the papers are concerned to find ways of preserving what has been found valuable in the universities as we have known them while facing up to the demands of mass higher education. Beloff, in particular, is trenchant in his defence of traditional virtues and he argues that 'we live in a society whose values are increasingly collectivist and egalitarian; a university must by its nature be individualistic and elitist'. However, Bell, who relishes the pejorative use of the word 'elitist', is frankly pleased to find that the 'traditional' role of the universities is, in fact, of comparatively recent origin and, he might have added, not all that distinctive, for Grant, in his comparative study, shows that while 'in every system there is some general concept of "university", it is often ill-defined and difficult to distinguish unambiguously from other kinds of post-secondary education'. One of the ways of coping with increasing numbers that is quite credibly argued—independently by Young and Thompson—is the introduction of tertiary colleges (16/17–19/20) to come between school and university, and Wolfe arrives at a structurally similar solution by proposing that some universities should 'concentrate' on postgraduate studies.

Several of the contributors, notably Barnes, trace the gradual erosion of university autonomy to the point where now the universities are almost totally dependent on the support and goodwill of the Government itself. It is healthy, therefore, but nonetheless disturbing, to have John Vaizey's expert exposure of the deficiencies and biases of the Government's projections of places and costs as set out in Education Planning Paper (No. 2). And we still await Planning Paper No. 3!

Much has changed in the four years since these papers were first presented. There has been the James Report and the White Paper, but, more important, there has been a change in mood. In the face of all the changes, the increasingly severe economic constraints, and falling student demand, the expansiveness and optimism of the sixties has given way to defensive uncertainty. To some extent this was foreshadowed in *Present and Future in Higher Education*, but its main interest now is in how a group of distinguished academics were thinking at the beginning of the decade. Some of the problems they were seeking to tackle are with us still, and in even greater degree.

ALAN SMITHERS

Advances in Educational Psychology

edited by Mia Kellmer Pringle and Ved P. Varma
University of London Press, 1974, 230 pp.,
Hardback £4.70

It depends what you mean by *educational psychology* and *advances*. The collection has two aims according to the authors: first it is a tribute to Professor W. D. Wall from colleagues and former students. Second, it '. . . presents a picture of research and thinking in the three fields of educational psychology in which Professor Wall has played a major and often innovatory role, namely learning and educational achievement; exceptional children; and adolescence'. The book treats of these subjects in three parts each part comprising articles by different contributors. Part one has papers on how children learn new skills, the purpose of futurologic studies in education and the affective consequences of school achievement. Part two has papers on the epidemiology of handicap, maladjustment, remedial education, brain damage, mobility training for blind children, the education of deaf children, gifted children and their education, and advances in training educational psychologists. Part three comprises papers on the psychosexual development of adolescents, the intellectual aspects of adolescence, social aspects of adolescence and physical aspects of adolescence.

One cannot doubt that the collection achieves the first of the editors' aims. It is less clear that it achieves their second. The problem I had is illustrated by the need I felt to list the various contributions to convey to readers the content of the book. *Educational psychology* is a term with catholic connotations but the selection of papers made seems to me to belie the title in two ways. On the one hand, as the editors say in their preface, it treats of the areas of educational psychology that Professor Wall has worked in, which implies a more restricted coverage than the title suggests (*Some Advances in Educational Psychology?*), and on the other, the term is stretched beyond breaking point by the paper on futurology which has nothing to do with educational psychology as I have seen the term used. (*Some Advances in Educational Psychology and Other Subjects?*) On the subject of *advances* some papers do, indeed, describe new developments in their varied fields of discussion, but others keep us firmly in the past with little sign of the new horizons which I hoped for when I opened the book. Perhaps the most appropriate title would have been: *Studies in Some Aspects of Educational Psychology and Other Subjects*. If this is your field of interest this may be the book for you.

When I reviewed the first volume in this series (*J. Child Pschol. Psychiat.*, 13, 1972) I expressed the hope that we might see a fuller treatment of school learning in a future publication since it was hardly mentioned in that book. The first contribution to the present volume does give promise of a different orientation especially in the early pages when, in the chapter on how children learn new skills, we read that the chapter is to deal with activities deliberately chosen by the teacher to bring about learning in children. However, the chapter is disappointing in its treatment of the subject. Not only is there nothing that could be remotely regarded as a discussion of *advances* but there is a musty smell about the pages. Readers may get some idea of the level by the suggestion that one can illustrate the passage of time in history by a succession of historical characters from different epochs as a kind of frieze along the wall of the primary classroom. In the same section of the book Husen writes on futurological studies under the same overall heading. This chapter contains some rather unexceptional remarks about likely future needs and developments in education. Some comments, however, remind us painfully of the perils of confident prediction. One is the assumption of continued economic growth and unlimited consumer horizons; another is the 'enrolement explosion' i.e. the increased proportion of cohorts going in for higher education. But, as I have already said, it is difficult to discern how this chapter relates to the title of the book or the heading of the section, 'Learning and achievements'. Bloom's chapter on the affective consequences of school achievement is more related to the theme. His discussion on self-concept, attitudes, motivation, and mental health in relation to success and failure in learning are of interest but little of new.

The section on the needs of exceptional children (taking *exceptional* to include those handicapped by high IQ as well as other disabilities) is more of a piece, although the editors manage to squeeze a rogue in here with a piece on the training of clinical educational psychologists. This particular chapter is in part an historical survey, in part a review of present practice and some speculation about the future. There is very little new about the training of educational psychologists and no account of 'advances'. Among other contributions in this section, there is a lively and stimulating review of the literature in the field of brain damage by Gardner, an interesting and informative review of the literature in the field of attitudes and methods in remedial education by Gulliford, and a useful review of the literature in the field of maladjustment by Chazan. Tizard provides a summary of the Isle of Wight survey and draws attention to the significance of the findings for the school psychological service. These four chapters may be reasonably taken to provide syntheses of current knowledge in the field and to point to the future. It is difficult to say as much for the rest of the chapters in the section. Burt's chapter on the gifted child has a high historical weighting and his espousal of the simplistic correlating of N.ach and kilowatt/hours per capita as criterial in assessing the progress of mankind, has an irony reinforced by his concluding sentence, when he comments that giftedness is not in itself enough, but that ideals and incentives are of almost equal moment. Robb's chapter on the education of gifted children is almost platitudinous in places. He makes much play of the need to train children in the skills of reference (the use of dictionaries, encyclopedia and the like) and makes the incredible confession that he has never seen this taught in any school he has visited. He writes that he believes that if '. . . a junior school could borrow a set of encyclopedia for three weeks only' for the teacher to train the children in the skills of reference, the pay off would be enormous. What chance have we of learning about advances from an author who has still to catch up with the present? The remaining two chapters in this section deal with provision for education and training of the blind and the deaf. Only tangentially do they relate to educational psychology and mainly describe provision for training.

Section three of the book has the rubric of *adolescence*. The main orientation is developmental, another example of the elasticity of the category *educational psychology*. Lovell's chapter on conceptual thinking at adolescence is probably the closest thing to a psychological exposition. Hemming's remarks on emotional and moral aspects of adolescence are interesting and related to education. Tanner is informative on physical development but there is little directly related to the psychology of adolescence. Mays has a chapter on the sociology of adolescence; a sympathetic and perceptive view of youth but again little to do with psychology. The paper by Osterrieth on the drawing of the human figure in relation to psychosexual development no doubt has relevance to problems of adolescent learning, but the layers of interpretation would probably be somewhat difficult to penetrate. It is, at any rate, the report of a recent investigation and represents, I suppose, an advance.

All in all this collection was for me a disappointment. It lacks unity and does not bear out the promise of the title. It may be that the editors set themselves an impossible task by trying to work to the title and at the same time to cover the diverse interests of a man who had worked in many fields. It may be that there are no 'advances' to chronicle in the field of educational psychology.

EDGAR STONES

Educational Research in Britain: Vol. 3

edited by H. J. Butcher and H. B. Pont
University of London Press, 1974, 420 pp.,
Hardback £4.60

This third volume continues the tradition set by its predecessors and contains 17 commissioned articles reviewing a wide range of educational research. The topics have been chosen to complement those in the earlier volumes so that there is no unnecessary overlap in the work described. An example of this efficient editorial procedure is provided by the first two articles: in these the development and funding of British Research as a whole is considered, whereas the previous volumes began with descriptions of the research being carried out by specific institutions, especially the NFER.

However, some of the suggestions made in *Educational Research* in 1971 (February) following the publication of Volume 2, also apply to Volume 3. A work of this kind is useful and essential for the researcher who wishes to gain an accurate impression of what has been done in areas outside his speciality: it is less useful as a starting point for new research, partly because the references, although they are fuller than in the previous volumes, stop at about 1971, and partly because the needs of a wider audience have had to be considered. So, while recognizing the high quality of the individual articles and the usefulness of the volume to a few, one can only amplify the proposals of the review of 1971: teachers and administrators want a paperback containing short but comprehensive articles linked by editorial comment; researchers want annotated bibliographies which can be added to from time to time.

D. DUCKWORTH

The Rise of the Schooled Society

by David Wardle
Routledge & Kegan Paul, 1974, 182 pp.,
Hardback £3·25

This very readable 'history of formal schooling in England' takes its cue—but not its conclusions—from the advocates of deschooling. Before contemplating abandoning our formal education system, Dr. Wardle argues, we need to review how it arose in the first place, as a means of getting to grips with just what it is. He then takes us on a rapid but vivid tour through the various phases of the development of the system.

Attention is given primarily to the climate of opinion from decade to decade, as evidenced in writings from government and other circles, and in the development of the major social forces and movements, such as technology and nationalism, as they affected educational policies. What emerges clearly is that the growth of the school system as we know it was the result of a recurrent *coincidence* rather than convergence of right and left wing policies: the various tendencies supported the development of schooling for very different reasons. The requirements of industrialists for a (semi-) educated workforce and of nationalists and militarists for more efficient troops were matched by the wish of working class people for education as an element of the good life, and by the belief of liberal reformists that universal education was a prerequisite for a functioning democracy.

Despite the fact that many of the original motivations have become much more muted in recent decades, this historical perspective throws a good deal of light on the state of tension in which schools find themselves today, for the structure of the education system and many of the assumptions operative in it can be traced to these early debates. The major weakness of the deschoolers' approach is seen, in Dr. Wardle's last chapter, to be their view of the education system as a monolith, the exclusive instrument of a manipulative ruling class. While there is plenty of historical justification for this view, and blatant examples of it are quoted here, there is also historical justification for seeing the education system as a testimony to liberal and socialist ideals. 'For every nineteenth century writer who favoured education as a means of saving society from upheaval we can quote another who hoped, or feared, that it would have precisely the opposite effect. . . . Schooling is not something which has been forced upon an unwilling public by an all-powerful establishment, but something which has been granted, rather grudgingly at times, as a result of overwhelming demand.'

None of this can save contemporary schooling from those indictments which really hurt, but Dr. Wardle's well-marshalled evidence suggests that our future policies for education are unlikely to be effective, and could be disastrous, if, like the deschoolers, we see a single-minded reactionary edifice where there is in fact a confused, ambivalent institution encompassing a variety of contending motivations.

G. CHANAN

The Teacher in a Changing Society

edited by John D. Turner and J. Rushton,
Manchester University Press, 1974, 98 pp.
Hardback, £1·50

This collection is the edited version of a series of lectures by five professors of education and one director of a university institute given at Manchester School of Education early in 1973. Represented are Professors Brown and Eggleston of Keele, Professor Musgrove and Turner of Manchester, Professor Ross of Lancaster and E. L. Elvin, Director of the London Institute. The title of the collection is a version of that given to the original lecture series so it is probably unintentional that it should sound so portentous. Certainly the portent and the line-up are not borne out by the contents, which range from the unexciting to the platitudinous. No blame attaches to the plan of the lecture series, for every effort is made to maintain a thread of continuity. It is simply that none of these eminent men is able, in the confines of a standard lecture, to give us any inkling of just what, precisely, this 'change' consists in. The conventional wisdom of 'changing society' is repeated without any real attempt at serious analysis, let alone critique. S. J. Eggleston tells us, with the air of a man uttering dangerous sedition, that the de-schoolers have challenged the current definitions of the role of the school. His lecture deals with 'three steps in the conceptualization of the role of the school', steps which—handily—are representable by a diagram. But nowhere does he question the idea—shared by the daring de-schoolers—that schools exist to deal in a commodity called 'knowledge' that is somehow ready-defined out there in that 'changing society'. It is as though Michael F. D. Young and the Goldsmith ethnomethodologists had never existed, and the sociology of education was still exclusively a matter of structures, roles and mechanisms.

Elvin treats us to a touching glimpse of childhood as the son of a socialist at a local council school—an early experience which no doubt helps to explain his fondness for that 'avoidance of extremes' for which he quotes George Steiner as believing to be Britain's special contribution to the European Economic Community. He is comforting about the moralism of 'our present society' compared with other eras. He is liberal in attributing equal morality to competing value systems: yet, in the middle of this affirmation of relativism he is able to say 'To have a child you know will have no father is to commit wanton injustice on the helpless'. In other words, like the other contributors, he is unable to find any validation for the ethical imperatives of the 'changing society' outside his own personal view. Indeed he goes further: he advises against any too definite a commitment about the future direction of society because 'we do not know what the society will be'.

And this is the problem of the whole series of lectures: none of these professors of education is prepared to suggest that 'society' is what people *make* it—and that the role of teachers may well be crucial in this respect.

Instead, A. M. Ross in outlining probable changes in teaching is constrained to 'play the game of futurology' and speculate upon the 'problem-centred' curriculum and other wonders. Teaching, he thinks, need not be thought of only in terms of society but can be justified in terms of 'self-expression, self-development and self-realisation'. He does not explain how the self can be realised *except* in terms of society, or what 'society' may be if it isn't individual people in social relationships with one another.

On the question of preparing would-be teachers, Musgrove talks of training guides for educational Cook's tours—whether facetiously or not is unclear—and ends up with the discovery that drop-outs in Manchester know about Hegel, while Turner indulges in speculation about White Papers and government intentions in a fairly predictable way, ending with the unexceptionable but hardly arresting observation that 'teacher education in the future must become more professional and more directly related to the work which the teacher undertakes in the classroom'. The most interesting—perhaps because least familiar—lecture is that of G. N. Brown on the needs of Africa but even he has an air of fatalism which sees the solution in terms of aid rather than the liberation of human consciousness. Perhaps expecting professors of education to talk compellingly about learning and learners is like expecting a synod of bishops to be interesting about sin: they are at the same time too involved and too out of touch.

DOUGLAS HOLLY

What School is For

by Gabriel Chanan and Linda Gilchrist
Methuen, 1974, 130 pp., 65p.

This book is well worth reading, and comes near to being good: not because its main arguments are (even when intelligible) satisfactory, but because the authors have a genuine feeling for the needs of large numbers of ('working-class') pupils and their relationship to our existing school system. Much of what they say about this relationship, which on any account is unsatisfactory, seems to me valuable and important: particularly when they concentrate on detail and authentic data rather than on generalization. Further, they are prepared to be critical of gurus like McLuhan (p. 8) and Marx (p. 10), and sufficiently sophisticated to avoid at least the more obvious sociological fantasies.

Like nearly all those writing primarily from a sociological viewpoint, however, the authors are extremely reluctant to give any clear account of what it is to be an *educated* person: that is, of what *particular* benefits or 'goods' are to be associated with 'education' (rather than, say, 'happiness', 'social progress', or (p. 127) 'an emerging culture of universal humanitarianism'). They are hence extremely reluctant to adopt a 'condemnatory attitude' (p. 62) to *any* 'culture' or state of ignorance or uneducatedness, or to regard *any* people as 'barbarians outside the gates' (p. 66, quoted from Peters). Yet any clear criteria will entail making judgements about people as 'uneducated'. In normal English, such criteria would roughly follow Peters' account: we should use words like 'ignorant', 'uncultured', 'barbarian', and so on: and we should allow that a person can be well-educated even if he had no particular interest in 'social problems' or 'the values by which we must live in the last quarter of the twentieth century' (p. 129), or even if he was a Fascist.

The authors seem to prefer other criteria; but I am frankly unsure (even after two readings) whether they really believe in any criteria which are not culture-bound or time-bound. It is one thing to accept 'working-class culture' because it is *there;* quite another to endorse it as *good.* Unless these criteria are clearly stated, one can hardly start; and if they were clearly stated, notions like social class and sub-culture would seem less relevant. To learn mathematics, how to read, how to do history, how to recognize and appreciate great works of art, and so on, must surely have dimensions which are in no sense social or sociological. The *point* of 'education as intervention, not initiation' (p. 63 ff.) is wholly obscure unless we know what counts as *success.*

This makes philosophical enterprises (such as Hirst's 'forms of thought', to which the authors do not refer) highly relevant to books like this one, and it is a great pity that the kinds of arguments almost any philosopher of education would advance are not here made clear or discussed—all the more because, as the authors imply, such arguments are often over-stated and need to be seen in a social context. This is, I suppose, just another example of the lack of proper communication between disciplines which vitiates so much educational theory and practice; but it does not detract from the other merits of an interesting book.

JOHN WILSON

New Movements in the Study and Teaching of Biology

edited by Cyril Selmes
Maurice Temple Smith, 1974, 399 pp.,
Hardback £3·00

'*The optimum environment of any organism is the one in which it evolved*' says Iltis (p. 296), but as the dinasaurs found out to their cost, environments change and organisms must adapt or die. The teaching of biology is no exception to this basic principle and laborious efforts have been made to change its image from that of traditional 'O' level biology, redolent of polished laboratories and formalin (with nature study for the masses); now it is the masses who have a watered down traditional course, often called CSE Mode 1, while the élite enjoy doing rather than hearing, studying the live rather than the dead, with notable exceptions in each case. This book shows clearly, however, that as yet there has been little attempt to define precisely why exactly it is taught as it is. The scientists and teachers who have contributed chapters seem mostly to base their science on Poppers' ideas, rationally testing their hypotheses to destruction, but their educational explorations seem more aligned to Kuhn's proposed mechanism (outlined in his book *The Structure of Scientific Revolutions*, Chicago UP, 1972), that of filling in the gaps of their knowledge according to the paradigm currently in favour. The acceptance of a current paradigm means that there is very little debate about fundamentals, and this lack of debate is compounded by the poor brand image which educational philosophy currently holds among teachers in general and science teachers in particular.

This book is full of implicit value judgements, for example (p. 4) 'pupils need to "get the feel of being scientists"' quoted from the Newsom report in the introduction—why should they? And if so, what sort of scientists? Porton Downers? Serious jokers like Crick and Watson? The variations are endless! None of these and similar underlying questions are answered; recipes are given, much dogma solemnly repeated. Do pupils learn much besides techniques from the kind of practical work outlined in Hornsey's chapter on radioisotopes (p. 188)? Many will merely follow the recipe—but the perceptual and intellectual problems involved in interpreting results are more to the point. Possibly a debate should be initiated as to whether radioactivity in school should, like electron microscopy, be high on the interpretive emphasis and low on the technology. Unkind critics have suggested that high technology in schools is merely an ego-trip for the teacher, a socially acceptable substitute for playing with trains; while rejecting such hyperbolé, Dunham's chapter (p. 239) is welcome in that he stresses the importance of the attitudes and motivation of the teacher as well as those of the pupils. Those promoting new ideas in science education must take more notice of consumer opinion than they did 20 years ago a relatively violent environmental change, unnoticed by many contributors to this book whose feelings are summed up by Stonemans somewhat sanguine remark (p. 268) that in education 'standards of success all exist within the system, independent of the outside world'. Pupils, parents, taxpayers and employers alike demand a rationale and in the absence of a sound philosophic basis are likely to assess the value of change on a time/cost/benefit basis. Recent newspaper headlines have demonstrated clearly that both civil servants and dumper truck drivers would have benefitted by education concerned with the principles of ecology and conservation, but it is doubtful whether Dowdeswells' neatly designed objectives (p. 114) would get a high assessment rating either on the criteria of cost/benefit or those of significance and relevance; in times of financial stringency these are the criteria which will be applied, if there are no others on offer.

New movements in the study of biology? Not particularly, although at least teachers are not proffered death camps for Drosophila any more, but Biston Betularia lives on! Much of the biological fact is unsystematically presented in an uncommunicative style which is likely to motivate few to read and less to act. Geddes' chapter on biophysics is an exception and communicates the excitement the topic holds for him, but some practising teachers would quarrel with his conclusion (p. 54) that 'biophysics *per se* cannot really be taught in schools. Without some background of solid state physics it will probably fail to make any impact. . . '. Come, come, it is possible to teach at a simple level of complexity without being simplistic! Teachers have engendered not only excitement in this topic but motivation to further, more complex work from the most unlikely pupils. But motivation to use mathematics as an appropriate symbolic language rather than a feeling of failure best avoided, does not spring from Dudley's chapter on page 168. Meaning is often tied to the sociocultural context of the communication, and while giraffes, elephants and leaves may provide elegant examples, only those with particularly irritating guilt feelings about needing to incorporate more maths in biology teaching are likely to rise to such flies. There is a growing feeling (emerging paradigm?) in schools that biology has less affinity with maths and the physical sciences than it does with English materials, value judgements and teaching methods. Fox's chapter 'Science for all' and Jenks' 'Personal relationships' both show this clearly and emphasise the kind of biology which is relevant for the majority of the school population and essential for those moving on to university. Communications experts can be seen at work in Lister's 'Assessment' and Reid and Booth's 'Independent learning'; these two chapters should be required reading for all teachers whether in practice or in training.

Consumer opinion is given its rightful place in Selmes final section, entitled 'The subversive approach', wherein strategies to allow students to plan their own courses are mooted; if this is subversive, then some London Medical Schools and not a few CSE Mode 3's stand accused. And the consequences could be grave! When a boy's school gave its pupils their own choice of route through the eight themes of the Nuffield Secondary Science project, they chose the most expensive route possible, while at the other end of the scale, there was once a college where everyone spent so much time discussing what the course should be that there wasn't much time left for the course itself.

Dorothy Dallas

Geography and Younger Children

by Eric J. Barker

University of London Press Ltd., 1974, 133 pp.,
Paperback £1·54, Boards edition £2·90

The publisher describes this book as in a sense a memorial to Mr Barker who died shortly after its completion. It brings together the reflections of the author on the teaching of geography to younger children after extensive experience as a teacher, college of education lecturer and author of textbooks.

The book is divided into three parts which correspond roughly to aims, methods and content but they are not strongly differentiated. The section on elements of geographical teaching, although the longest, does not contain enough examples of classroom practice to form a guide for young teachers. Instead, the whole book should be considered as an extended essay on the teaching of geography to juniors and, as such, will be of most use to teachers of some experience. Viewed in this way the first section of the book is especially valuable, for here are exposed many of the problems which are endemic to curriculum development in junior and middle schools—the relation of subject discipline to general skills, the problem of underlying values, criteria for selection of material and so on. Mr Barker has written modestly of his own views on these issues and one wishes that the book might have given a clearer exposition of the provisional solutions which he must have found. Only in Chapters 4 and 5 which deal mainly with map work can one distinguish a progression of skills associated with geography. The book seems most convincing, as in Chapter 7, when the author, temporarily free from the constraints of a discipline, writes as a teacher of junior rather than as a teacher of geography. There is a spirited defence of the use of junior school texts which should be read in conjunction with an earlier remark that 'the criterion that can be applied to judge the success of a period is easily stated. It is simply a question of the degree to which the initial objective set by the teacher was in fact achieved.'

It is a sign of the times that many, perhaps a majority of teachers would not regard this comment as a truism. Mr Barker's book will have achieved a great deal if it has persuaded some that their classroom planning should perhaps resemble the behaviour of an enthusiastic punt captain more closely than that of a fog bound ocean racer.

A. J. Heaman

Physical Education in Secondary Schools

by J. E. Kane

Macmillan Education, 1974, 118 pp.,
Hardback £3·00

John Kane has produced an informative monograph, whose declared intentions were to clarify the place of Physical Education in the curricula of secondary schools and to provide a basis for a curriculum development project. There is no doubt that *Physical Education in Secondary Schools* succeeds in providing detailed descriptive information which usefully outlines for P.E. teachers the whole area in which they work. However, the report does not indicate the direction in which any curriculum development could proceed.

The results of the investigation on which the report is based are clearly and systematically presented, and the text is well supported by 101 tables and seven figures. The research data were collected by a mailed questionnaire which had an 82 per cent response rate. The face validity of the questionnaire is high, and Dr. Kane and his assistant, Miss June Layson, are to be congratulated on their achievement in producing so comprehensive an instrument which 'had to be discussed, constructed, revised, piloted, analysed, adapted, printed and despatched in about 10 weeks'. Nonetheless, one would have liked to have been given some evidence of the reliability of the sections of the questionnaire concerned with Pupil Effects, Influencing Factors and Teacher Characteristics, where the judgment of the responder was required. Mention is made in the introduction of 'an intensive search through a considerable range of literature', but unfortunately no references are given.

Several interesting points arise. On page 23 Dr. Kane asks the question, 'Will the part-time coach or expert in a specialist field become increasingly prominent in our larger schools?'. This raises the question of the training of such coaches for coaching in schools. Is this not a role for the advanced physical education courses in colleges of education in conjunction with the national bodies of sport?

The picture of the ideal PE teacher emerging from the teachers' assessment of the importance of certain teacher characteristics is that of a hard working, honest, knowledgeable, inspiring communicator of ideas. Perhaps influenced by the archetypal Games Master in the film 'Kes', they gave a low ranking to 'extraverted personality', but on the other hand giving low rankings to cultural knowledge and academic achievement does not help the image.

In the ranking of objectives, one difference between the men and women respondents deserves note, i.e. that physical education as a preparation for the enjoyment of leisure time is ranked highest out of nine by the men, but only fourth by the women. The national growth in the popularity of leisure pursuits in the past decade, reflected in the spread of sports centres of various kinds, leads one to ask whether this aspect of physical education has, perhaps, been a little neglected in some teacher education courses.

The questionnaire on objectives does not bring out fully the attitudes of the sample of PE teachers to their role in the promotion of optimal, all round physical development. The high rank given to 'motor skills' was predictable, given its wide terms of reference, but the low rank given to 'organic development', by younger teachers particularly, might be an indication of some neglect of one of the vital purposes of physical activities in school. Despite the growing interest in leisure pursuits, these are still confined to a minority of the population; most secondary school children are spectators rather than participators outside school hours, and it would therefore seem to be an important necessary objective of a comprehensive physical education programme to ensure that pupils get enough physical exercise (or movement experience!).

K. Challinor

Spelling: Task and Learner
edited by Barrie Wade and Klaus Wedell
Occasional Publications Number Five,
University of Birmingham, 1974, 68 pp., 85p.

It is an interesting reflection on the attitudes of our society that whereas poor standards of handwriting—sometimes to the point of illegibility—are not only tolerated, but occasionally proclaimed a virtue, irregularities and idiosyncracies in spelling are identified with being uneducated. It was not always so, of course: before the days of Dr. Johnson individualised spelling was as much an indicator of a person's character as some people would have that handwriting is today. Certainly a case can be made for allowing personal freedom in spelling and handwriting skills, while retaining a standard form for printing.

Unfortunately, for the time being, at least, we have to accept facts as they are, and poor spelling as much as poor reading is regarded by most people as a stigma of the illiteracy prevalent in our society. So felt those who attended a Study Group on Spelling Disability in Children held at the School of Education in the University of Birmingham in July 1973, and this publication contains the papers and a summary of the discussions held.

For those who share the concern about bad spelling—and perhaps for many who do not—this book contains much to interest them. The papers range from a linguist's attempt to annunciate the lexical rules of our writing system—generally considered far too difficult to be of value in teaching children—through a neuropsychologists approach to the aetiology of spelling disability and a study of spelling errors, to the more practical aspects of remediation. More space might have been devoted to the discussions, but, as with most conferences of this kind, their usefulness was probably restricted to the participants. Although apparently dealt with at some length at the meeting, no satisfactory answer was given to the question of how children learn to spell, and the fact that, for many children spelling difficulties arise because they have to learn the irregularities of our orthography at the same time as they learn to read, was not even mentioned. There was unanimous agreement, however, that more spelling must be taught in schools, and that teachers need more training in how to teach it.

DOUGLAS PIDGEON

RESEARCH IN EDUCATION

May 1975 *Number 13*

The impact of social science on flows from school to university: D. HUTCHINSON and G. M. LITTLEJOHN

Objective test analysis: some criteria for item selection: J. WILMUT

The satisfaction of university staff and their teaching: RICHARD STARTUP, MICHAEL GRUNEBERG and PATRICK TAPSFIELD

Syllabus-bound and syllabus-free orientations in a technological university: ALAN SMITHERS, PAT SAYLES and ALAN JOSEPHS

Some thoughts on norm-professional and criterion-referenced measures: T. J. HARVEY

Published twice-yearly, May and November
£4.50 per annum, single copies £2.75

MANCHESTER UNIVERSITY PRESS
Oxford Road, Manchester M13 9PL

NFER

COMPUTER SERVICE

The National Foundation for Educational Research has now installed its own computer—an IBM 1130.

We anticipate a certain amount of spare time in the next year or so, and are prepared to make the computer available to Educational Research Workers at favourable rates.

Card-punching will also be undertaken, at competitive rates.

Details

A comprehensive range of programs is available. These include regression analysis, factor analysis (up to 100 variables), a general tabulation program, χ^2 analysis, and analysis of variance. In addition, there is a wide range of data processing programs for those researchers who need to handle a large amount of data.

Rates

Computer: £25 per hour.

Data Processing: £30 per 1,000 cards. Punched and Verified.

Specimen Costs:
Regression analysis of 8 variables on 40 subjects—£6: 15 tables from 2,600 subjects—£12.

All charges subject to 8% VAT.

Telephone Slough 28161, or write to
A. N. James,
NFER, The Mere, Upton Park, Slough, Bucks.

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £2·75 post free. Overseas (excluding USA and Canada*): £3·20 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: 95p post free. Overseas*: £1·10 post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA and Canada: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Inquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Berks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the text, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or
2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© National Foundation for Educational Research in England and Wales, 1975

Printed by King, Thorne & Stace Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

25 SEP 1975

In this issue

Stimulus material in O-level history
Attitude measurement: a critique
Teacher life changes and student dropouts

Acc [illegible]
Docket [illegible]
[illegible]
24/1/76

Published for the
National Foundation
for Educational Research
in England and Wales
by NFER Publishing Company Ltd

EDUCATIONAL RESEARCH

Volume 17 · Number 3
June 1975

Contents

Standards of reading: a critical review of some recent studies *E. Burke* 163

Measurement of human values: an appraisal of the work of Milton Rokeach
T. M. Kitwood and A. G. Smithers 175

Can pre-school children form concepts? *R. M. Povey and E. Hill* 180

Primary French in the balance *C. Burstall* 193

Personality, ability and achievement in primary school children *K. Orme* 199

The reliability of teachers' judgement of the abilities being tested by multiple choice items
R. W. Fairbrother 202

Alternatives to the traditional sixth form: a preliminary report *L. Orr and J. Dean* 211

Further education: an alternative route reviewed *G. M. Young and M. Macdonald* 221

The educational needs of the mentally handicapped *J. M. Hughes* 228

Book Reviews 234

Diary No 942 (2)
Date 25/7/75
[illegible]

EDUCATIONAL RESEARCH

A review for teachers and all concerned with progress in education

The journal of the National Foundation for Educational Research in England and Wales

Index to Volume 17 (1974-75)

page

PHYSICAL EDUCATION IN SECONDARY SCHOOLS *by J. E. Kane* 158

SPELLING: TASK AND LEARNER *edited by B. Wade and K. Wedell* 159

RUNNING A SCHOOL *by C. H. Barry and F. Tye* 234

EDUCATION OR DOMINATION *edited by D. Holly* 234

OBEDIENCE TO AUTHORITY *by S. Milgram* 236

PLAYING, WORKING, GROWING *by G. Ivimey* 237

SHAPING THE EDUCATION OF SLOW LEARNERS *by W. K. Brennan* 237

TEACHING LANGUAGES TO ADULTS FOR SPECIAL PURPOSES *CILT Reports and Papers II* ... 239

THE EDUCATION OF THE POOR: THE HISTORY OF A NATIONAL SCHOOL 1824–1974 *by P. Silver and H. Silver* 239

Printed by King, Thorne & Stace Ltd, School Road, Hove, Sussex, BN3 5JE

edited by
Alfred Yates
Clare Burstall
Gabriel Chanan and
Monica Taylor

EDUCATIONAL RESEARCH

Volume 17 Number 3
June 1975

Standards of Reading:

a Critical Review of Some Recent Studies

Elizabeth Burke, Colleges of Education Division, Faculty of Education, University of Manchester, and
D. G. Lewis, Department of Education, University of Manchester

Summary

Six studies of the reading standards of British school children are critically reviewed, special attention being paid to the statistical analysis and evaluation of results. No justification was found for a belief that standards are declining, and the implementation of further periodic, regional surveys is suggested.

Introduction

Standards of reading attainment in our schools are of central concern at the present time. The 'Committee of Inquiry into Reading and the Use of English' (the Bullock Committee) has recently reported, and large-scale researches financed by the Schools Council are at present in progress in the Universities of Manchester, Nottingham and Sheffield.* There is also the SSRC financed project into the teaching of reading in Scottish schools.†

In addition, a number of individual research studies are being undertaken in this area. The results that emerge may well suggest revision of our present methods of teaching reading, of the curriculum of the primary school and its time allocation to the development of reading skills, and of our training of teachers for remedial education.

Much of the evidence collated by the Bullock Committee is concerned with present practices in schools and with the opinions of teachers, administrators, reading specialists, child psychologists and others on what changes, if any, should be made. Surveys of existing practice and of informed opinion are of value. But of at least equal worth is the evidence of planned research. It is important too, not only that the implications of such research be understood but that its limitations be appreciated. A number of relevant research studies have been published in recent years. A critical review of those that seem of major importance may then be helpful. The studies selected for this review fall under one of two heads:

(i) studies that are essentially normative, being concerned with standards of attainment,

and

(ii) studies that intend to reveal what concomitant variables (or combination of such variables)—and, in particular, variables in the 'school' or 'teacher' categories—are associated with high standards.

Only studies that come under at least one of these heads are included, and again any aspects of the selected studies not falling under these heads are excluded. Each study is subjected to a critical analysis, such an analysis being necessarily selective and concentrating on those points of omission or of misrepresentation which are most likely to mislead.

A fundamental question to be answered at the outset of any research is the choice of criteria. How, in any investigation in this field, is reading to be assessed? It is desirable that any proposed measure be objective (i.e. uninfluenced by the observer or process of measurement), reliable (in that it is repeatable and not subject to short-term fluctuation) and valid (agreeing reasonably well with other superficially equally plausible measures). Many investigators choose a group test of reading comprehension. Here the test items may require the selection of a word, from a given set of options, to 'complete' an incomplete sentence. Again, at a more advanced level, one could aim at the comprehension of a larger unit such as a paragraph of connected sentences. Other measures could focus on the speed at which a child can read, or on accuracy of pronunciation, or on whether he reads with expression. We could also depart from the purely cognitive and attempt to assess the interest in reading which a child has acquired (by noting the proportion of his 'free' time he chooses to devote to reading possibly). The point here is not that an investigator should attempt to include all, or even most, kinds of possible criteria, but rather that he should be aware that the initial problem is one of selection, and that other criteria of success are possible, and that any single criterion study is necessarily limited from the outset. Indeed we could go further and argue for a multi-criterion designed research whenever possible, if only to probe the extent to which the pattern of results differs for each of the separate criteria (Lewis, 1972).

---

*Extending Beginning Reading (Southgate Booth); The Effective Use of Reading (Lunzer); Children's Reading Habits (Whitehead).

†Teaching of Reading in Scottish Schools—(Maxwell, Queen Margaret's College, Edinburgh).

A second basic issue—and one of equal importance to the choice of criteria—is one of sampling. This is because we are seldom concerned with the obtained results as ends in themselves. Rather we wish to generalize to other groups, to a (relatively large) population of which the group tested is representative. We have to concern ourselves then with such questions as whether the pupils tested in the study have been sampled directly, or whether the sampling was of more than one stage. Was it, for example, a two-stage or a three-stage process, 'areas,' 'schools within areas' and 'pupils within schools'? Again if 'schools' have been built into the design as a sampling unit, have certain categories of schools been excluded, for example, independent schools and schools for special categories of children such as the maladjusted? When the design has been decided it follows that every stage of the sampling process should be reflected in the subsequent analysis of results (Lindquist, 1956; 2-5; Lewis, 1968, 88-95) otherwise an inappropriate estimate of statistical error becomes inevitable.

A further point, not always appreciated, is that the selection of a sample inevitably implies a definition of a population. Here we have to consider two inter-related questions—whether the sample is likely to be representative of this population, and whether this population corresponds to any actual existing population of pupils or schools. These two questions correspond to the distinction between the internal and external validity of an experiment (Campbell, 1951; Campbell and Stanley, 1963). The former is essentially the control exercised over extraneous variables and the latter the extent to which we can generalize to practical situations. Ideally, designs should be strong in both. Given the limitations of time and finance however, we have usually to steer a middle course between the Scylla of excessive control and the Charybdis of rampaging statistical error. Considerations such as these should always precede our evaluation of the results.

In the present case we have chosen to focus on six major studies on reading standards. After the aims, criteria and sampling procedures of the selected studies have been described, the findings, as well as the interpretations of the researcher, are critically appraised. An overall evaluation is added. A concluding discussion seeks to bring together some of the main issues brought out in the separate studies. The paper is intended as a contribution to the continuing debate as to how best to ensure standards of literacy in our schools.

The Morris Study (1966)

(a) *Aims:* The two major aims of this study were (1) to compare the reading attainments of primary school children in Kent with those of children in the country as a whole, and (2) to relate these attainments, as well as preceding progress, to methods of teaching, type of school organization and the non-verbal ability of the pupils concerned.

(b) *Criteria:* Two tests of reading comprehension were used for both the aims stated above, namely the *Watts-Vernon* (WV) and *National Survey Form Six* (NS6) tests. Each had been used in previous surveys and was of the 'incomplete sentence' type. The WV test, originally designed for a national survey of reading standards in 1948, has 35 items and a time limit of 10 minutes, while the NS6 test, constructed in 1954, with 60 items and a time limit of 20 minutes, provided more scope for the abler pupils.

In addition, for the longitudinal study necessary to assess progress, *Watts' Sentence Reading Test* (SR1) was used together with other measures.

(c) *Sampling:* The sampling unit was that of 'schools', 52 schools being selected, all part of the sample taking part in an earlier 1954 survey (Morris, 1959) and having pupils who since January 1955 still had eight terms of primary schooling to come. Thirty nine schools were junior with infants and 13 junior without infants. Scores from 2,353 fourth-year pupils, presumably all in the schools concerned (absentees on the day of testing apart), were obtained.

For the study of reading progress, scores from 1,848 pupils from the same 52 schools were obtained over the period 1955-57. In addition 12 'extreme' schools—three 'good', three 'poor', three 'improving' and three 'deteriorating' in respect of their pupils' reading attainment—were selected for special study.

(d) *Results:* A comparison of the mean scores of the Kent children on both the WV and NS6 tests with the national means (the Kent mean being adjusted to allow for age differences) provided the county authority with the assurance that the superiority of their reading standards found in the 1954 survey was being maintained (and possibly even increased). In deriving national means some extrapolation was necessary, but the amount was not excessive. More important than such prestige comparisons however—the results of which should obviously be set alongside corresponding differences in the socioeconomic conditions of the homes—would seem to be the finding that the reading standards in the schools themselves were being maintained.

Of far more potential interest however are the results pertaining to the second aim of the study, namely the link with other concomitant variables, and for this purpose the schools were grouped on the basis of the commencing method of teaching reading, phonic or whole-word. The pattern of the SR1 mean scores suggested a better rate of improvement for the whole-word method, the difference being statistically significant however only after adjusting for school differences in non-verbal ability. Morris appears to argue (*op. cit.*, 32) that the lack of significance between the two school groups in respect of non-verbal ability precludes a finding in favour of either method. But if the analysis of covariance is to be legitimately applied, in removing *chance* differences in non-verbal ability between the groups, it is this lack of significance that is vital. The evidence should not be discounted for this reason. The findings of course relate to two hypothetical 'populations' of schools of which the actual groups compared can be considered random samples.

Morris also investigates the effect of the type of organization, junior with infant schools versus junior without infants, but this variable is largely confounded with that of methods. There are in fact only three schools in one cross-classification. Obviously little of relevance can be extracted from this situation.

Finally, as regards the special sample of 'extreme' schools, the explanations suggested for these conditions—staff changes, influence of the staff and headteacher, etc.—should be viewed with reserve in that the design suffers from a fundamental defect of omission, the lack of control group of 'non-extreme' schools. It is possible for instance that the influence of the headteacher would have been equally dominant here.

(e) *Evaluation:* The finding that during the period of the investigation reading standards in Kent primary schools were maintained is clearly acceptable. The study is less successful in realizing its other aim of linking success in reading with method of teaching and type of school organization. No conclusion can be made here.

The National Child Development Study

(Kellmer Pringle, Butler and Davie, 1966)

(a) *Aims:* The general aim of this study was that of compiling normative data about the educational, behavioural, emotional, physical and social development of a large group of British children. As part of this aim, norms for reading attainments were included. It was also intended to concentrate upon the incidence and associating conditions of handicaps.

(b) *Criteria:* The *Southgate Reading Test* (Southgate, 1962) was used as a measure of word recognition. For each item children have to select from a number of options the word which corresponds to a picture in the test booklet. The test has a low ceiling for seven-year olds, being unable to extend the above-average reader, and is more successful in differentiating among the backward.

As a second general measure of reading attainment the number of the reading primer then being used by each child was noted. While recognizing the variation in the level of difficulty between books in the various reading schemes, it was nevertheless considered that meaningful distinctions could be made on this basis.

(c) *Sampling:* The theoretically available sample consisted of the surviving children of the Perinatal Mortality Survey (Butler and Bonham, 1963), i.e. all children born in England, Wales and Scotland during the period 3rd-9th March, 1958. Only children living in England, and in respect of whom information was received by mid-August 1955, were included in the sample however. The authors report some evidence of bias in that these early returns had a slight under-representation from the clerical and semi-skilled manual occupational categories. Data from 7,985 seven-year-old children were analysed.

(d) *Results:* The statistical analysis is largely confined to the reporting of descriptive statistics though some hypothesis testing also appears. Thus the null hypothesis of no sex difference in mean score on the Southgate Test is rejected, the difference in favour of girls being significant at the one per cent level. This result is confirmed too from a study of the reading primers used. More boys were still on Book 1 (for non-readers) and also on Books 2 and 3 (where the reading was mainly mechanical, with help from the teacher required frequently), while more girls were on Book 4 (an indication of true reading ability). The percentages of non-readers were 11·2 for the boys and 5·9 for the girls, the combined figure being 9·8, while the percentages of poor readers (Books 2 and 3) were 42·0 for boys and 31·2 for girls, the combined figure being 37·3. By comparing these findings with corresponding results from an earlier survey of Kent children (Morris, 1959), it is tentatively concluded that the number of poor readers and non-readers transferring to junior classes dropped during the years 1956-65. Apart from the inherent unreliability of the criterion employed it is doubtful however if even a tentative conclusion based upon a comparison of two surveys, one national and the other of an untypical county, and separated by a time span of over ten years, can be justified.

Of the various concomitant variables included in the study the one of major educational importance is undoubtedly that of *length of schooling.* The research is advantageously designed to isolate this variable in contrast to many investigations where it is confounded with age. Two groups are isolated, one of 'early starters' aged from 4·6 to 4·11 on starting school in January 1963, the other of 'late starters', aged from 5·0 to 5·5 on starting school in the summer term of 1963. A comparison of the proportion of

Table 1: The proportion of good and other readers for early and late starting children in each father's occupational group (from Kellmer Pringle, Butler and Davie, 1966)*

Occupational Group	*Early (E) or Late (L) Start*	*Good Readers*	*Other Readers*	*Difference* $(a+d)-(b+c)$
Professional and technical	E	58·0 (a)	42·0 (b)	
	L	47·8 (c)	52·3 (d)	20·5
Other non-manual	E	61·9	38·1	
	L	51·2	48·8	21·4
Skilled manual	E	39·2	60·8	
	L	36·7	67·3	9·0
Partly-skilled manual	E	33·1	66·9	
	L	28·2	71·8	9·8
Unskilled manual	E	25·7	74·3	
	L	22·5	77·5	6·4

**op. cit.*, Table 61, 122.

'good' readers—scoring highly on the Southgate Test—in the two groups showed a clear and statistically significant advantage for the early starters. The authors also take account of *father's occupation*, and show that the advantage of early starting obtains for each occupational group. They claim too that the advantage is independent of occupational group (*op. cit.*, 124). This however is not so. The breakdown for each group is shown in Table 1, where the figures in the last column give the difference between the row differences ('early' minus 'late' starting) for the preceding two columns (i.e. for the 'good' and 'other' readers). In other words the last column assesses the *interaction* between early versus late starting and reading, and we see that the extent of the interaction differs from group to group. The children of professional and other non-manual parents who have the advantage of an early start maintain this advantage to an appreciable greater extent than do the children of parents of manual occupations. There is a strong indication in other words, of a second-order interaction, that between reading, early or late starting *and* occupational group. The educational implications are of obvious importance.

(e) *Evaluation:* The study especially when supplemented by the inclusion of the 'late returns' provides valuable baseline data for future surveys. On the other hand, an attempted link up with a previous survey is best ignored. It is unfortunate that in the analysis of the interrelations of variables a second-order interaction involving length of schooling and parental occupation has been overlooked.

The Inner London Education Authority Literacy Survey (1969)

(a) *Aims:* The declared aims of this survey include (i) discovering the extent of children's reading difficulties sufficiently early to take any necessary remedial action in the primary school, and (ii) comparing reading standards in London with those nationally.

(b) *Criteria:* A group test of reading comprehension devised by the NFER for their streaming project (Barker Lunn, 1970) was used. The test is of the sentence completion type and has norms derived from a sample of 3,000 children in each year group.

(c) *Sampling:* All the junior schools of the Authority were asked to give the test in October 1968 to all their pupils born between 2 September, 1959, and 1 September, 1960. Test scores from some 96 per cent of the schools were obtained. With pupils from special schools excluded the number of pupils involved was 31,308.

(d) *Results:* The mean standardized score of the sample was 94·4, and this being relative to a 'national' mean of 100 and standard deviation of 15 shows the London eight-year-olds to be on average some six months behind in reading comprehension. Again, defining, 'poor' readers as those with a score less than 80, some 17 per cent of the sample (20·9 per cent of the boys and 11·8 per cent of the girls) are poor readers. This compares with 8·6 per cent of children nationally. Two very reasonable comments are made here. It is pointed out that nearly 17 per cent of the sample are immigrants, and overall the reading standards of immigrant children are lower than those of non-immigrants, the proportion of poor readers among immigrants being in fact almost double that for the non-immigrants. It is also pointed out that the intelligence of the London eight-year-olds may well be below that of eight-year-olds in the country as a whole.

The report goes on to illustrate the influence of *parental occupation* by showing that the proportion of good readers decreases markedly as one proceeds from children of professional parents to those of lower occupational status. It deals similarly with such variables as the *cultural stimulus of the home* (as assessed by the teacher), *parental interest in the child's education* (also as assessed by the teacher), *family size*, and *birth order*, the procedure in each case being that of simply reporting the proportion of poor readers and also that of 'good readers'—good readers being arbitrarily defined as those with a score of 115 or more—in each of the various levels of the variables concerned. The presentation is commendably clear throughout. Nothing more ambitious is attempted.

(e) *Evaluation:* The survey provides valuable evidence of reading standards in London primary schools, the very high proportion of respondents being greatly to the credit of the teachers concerned. As comparisons with national standards depend upon the norms of one test compiled from a standardization sample of only 3,000 eight-year-olds (a so-called 'national' sample) something less less than complete assurance is advisable here.

The Clark Study (1970)

(a) *Aims:* The aims of this investigation are three-fold, being that of

(i) determining the general reading standards of primary-school children in the county of Dunbarton, Scotland;

(ii) making a study of all found to be backward readers (backwardness in this respect being defined as a reading quotient not exceeding 85); and

(iii) making a further, more intensive study of those backward readers who were average in intelligence.

(b) *Criteria: The Schonell Graded Word Reading Test* (Schonell and Schonell, 1960) was used to determine general reading standards. This is essentially a test of word recognition, the numbers of words correctly pronounced being converted to a reading age for accuracy in the reading of isolated words out of context. The *English Picture Vocabulary Test* (EPVT) (Brimer and Dunn, 1963) was used as a supplementary measure. This is intended to assess levels of listening vocabulary, and each item consists of four pictures, the child having to identify the one associated with the word spoken by the tester.

For the second of the aims of this study as stated above the *Southgate Reading Test* (Southgate, 1962) was used, and the *Neale Analysis of Reading Ability* (Neale, 1966) was administered to all selected for the more intensive study of reading backwardness in the final stage.

(c) *Sampling:* The sampling was in three stages according to those of the research. For the initial stage, all Dunbartonshire children whose date of birth fell between 1 April and 31 August, 1959, and who were still attending a school in that county in June 1966, were tested. By restricting the dates of birth in this way all the sample had completed two years of schooling at the time of testing. 791 boys and 753 girls were tested.

The stage two sample consisted of all the stage one sample whose (Schonell) reading quotient did not exceed 85 (apart from the very few who had left the county during the interviewing year). This sample of eight-year-olds consisted of 138 boys and 92 girls. The sample was selected by intelligence only in so far as intelligence was correlated with reading. The mean WISC IQ was 89·8.

For the stage three sample, selection in respect of WISC IQ was deliberate. An experimental group of 46 boys and 23 girls (now as nine-year-olds) was selected and all with at least one IQ not less than 90 (all three of full-scale, verbal and performance IQ being considered) and with a 'poor' Southgate Reading score (21 or less). A corresponding control group was selected on the same basis regarding IQ but with a 'good' Southgate Reading score (22 or more).

(d) *Results:* The mean standardized scores obtained by the stage one sample on the Schonell test, 106·8 for boys and 108·3 for girls, were both above the mean of the standardization sample (100). This however should not be taken as evidence of a rise in reading standards partly because the norms may now be out of date and partly because on the similarly standardized EPVT means nearer to 100 were obtained (and as a measure of comprehension this second test seems the more relevant). Equally, of course, no hypothesis of decline can be sustained. Mean reading scores for different levels of drawing and right-left differentiation are recorded—with no connection between the variables being apparent—though coefficients of correlation based on individual measures are not reported. But the finding of main interest is undoubtedly that some 15 per cent of all Dunbartonshire seven-year-olds are backward readers to the extent that their (Schonell) reading quotient does not exceed 85. It is these backward readers who comprise the stage two sample.

The author finds that as much as 44 per cent of the stage two sample have WISC IQs under 90, and that if the *verbal* subscale IQ of 90 is taken instead this percentage figure rises to 51. Clearly then IQ emerges here as the variable of major importance. Other variables may of course have an even greater educational significance and the search for them should be carried out painstakingly. The results too should be co-ordinated preferably by a regression analysis where the relative importance of these other variables as *accompanying* predictors could be ascertained. Clark's analysis however is insufficiently detailed, being confined to a consideration of each predictor in isolation. At the same time what has been done in the analysis of results from the stage one and stage two samples impresses as sound even if the maximum information has not been squeezed out of the data. It is when we turn to stage three that this verdict has to be revised.

Despite the dominant role clearly played by intelligence, Clark decided to restrict her further, more detailed study of backward readers to those 'within the normal range' cf intelligence. At the same time a second (control) group was selected, of similar intelligence but of higher reading attainment (though still part of the stage two sample of backward readers). The criterion for eligibility for 'normal' intelligence is a curious one. It is that *any one* of the three WISC IQs—full-scale, verbal and performance—be 90 or more. There are two objections to this. One is that the three IQs are not parallel measures of the same trait, or combinations of traits, and the other is that the one-in-three procedure capitalizes on chance errors of measurement—many whose 'true' IQ (of whatever variety) is below 90 qualify for 'normalcy' in this way, and this is particularly so when the additional criterion of a very low reading level is also applied (as it is for the experimental group). Intelligence in fact has not been controlled between the experimental and control groups, despite the application to both of the same selection procedure as is seen from Table 2.

Table 2: Mean WISC IQ of the experimental and control groups of backward readers of normal intelligence (from Clark, 1970)*

	Experimental Group	*Control Group*
Boys	94·4	98·6
Girls	91·4	92·0

**Op. cit.*, Table g, 137.

A further criticism is that even if the scores had been equalized in respect of IQ the fact that the experimental group is an extreme selection of backward readers means that in any subsequent testing its mean score on reading—or on any correlated measure—can be expected to regress upwards, towards the population mean, i.e. towards the mean score of all the backward readers if they had been included in the testing (see McNemar, 1962, 161-162). There is in other words, a built-in bias for the experimental group's results to be inflated. Actually, group comparisons do not seem to have been followed up in full, the author's main interest appearing to be in individual case studies. Nevertheless, the defect in design needs pointing out.

(e) *Evaluation:* The normative aims of the study have been pursued with commendable thoroughness and some useful information on several psychological variables has been gained. The intensive study is less successful. It is unfortunate that advice from a statistician was not obtained at the planning stage.

The Start and Wells Study (1972)

(a) *Aims:* The primary aim of the study is that of assessing the attainment in reading comprehension of English children at two age levels, those of 11 and 15 years, and so making possible comparisons with previous surveys and providing a baseline for the future. A further aim is that of making comparisons among sub-groups, such as boys and girls and pupils from different types of schools. This however is conceived as of secondary importance, and might well have been an afterthought.

(b) *Criteria:* Again the *Watts-Vernon* (WV) and *National Survey Form Six* (NS6) were used as measures of reading comprehension, the former having been used in national surveys since 1948 and the latter since 1955.

(c) *Sampling:* The sampling was performed in two stages—schools and pupils, and that of schools was stratified both by *school type* (the strata being grammar, modern, technical, comprehensive and direct grant for the senior sample, and junior without infants and junior with infants for the junior sample) and *school size* (large, medium, and small). The actual selection was made randomly by computer from lists supplied by the Department of Education and Science. Within each of the co-operating schools, pupils were selected on the basis of their date of birth in any month. Two equivalent samples of junior schools were selected, one for each of the tests, while at the senior level only one sample was selected, the schools being asked to administer the WV to half their pupil sample and the NS6 to the other.

The proportion of schools declining to co-operate was disturbingly high, only 73 per cent of the primary schools agreeing to take part. Again, due largely to a protracted postal strike, only 53 per cent of the secondary schools were able to return data. Further data relating to the WV test was lost (some 9 per cent of the design sample) through an office mishap. The final yield was 73 WV schools (1,660 pupils) and 69 NS6 schools (1,470 pupils) at the junior level, and 135 schools (1,844 pupils) for the WV test and 161 schools (2,194 pupils) for the NS6 test at the senior.

(d) *Results:* The test mean scores for the present and previous national surveys were age corrected—to remove the effect of (slight) age differences between the various year groups—and for the junior level these were found to be as shown in Table 3. For neither test is the difference between the 1970 mean and the mean immediately preceding significant at the five per cent level, though for the WV test the difference only narrowly fails to attain significance at this level. Start and Wells point out that if the means had been corrected to the mid-point of the mean ages for the 1964 and 1970 surveys, the differences would then (just) become significant at five per cent. Numerically the difference corresponds to a reading age of 4·38 months. For the more reliable NS6 test the difference between the last two mean scores corresponds to a difference of only 0·24 months. The mean scores for the surveys at the senior level are as shown in Table 4. Both the slight rise in the 1971 NS6 mean and the slight drop in the 1971 WV mean are statistically insignificant.

Table 3: Mean scores for pupils aged 11 years 0 months (from Start and Wells, 1972)

A The Watts-Vernon (WV) Test
(Maintained schools only)

Date of Survey	1948	1952	1956	1964	1970
Mean Score	11·59	12·42	13·30	15·00	14·19

B The National Survey Form Six (NS6) Test
(Maintained schools only)

Date of Survey	1955	1960	1970
Mean Score	28·77*	29·48	29·38

*Estimated from the obtained mean applying to England and Wales.

Table 4: Mean scores for pupils aged 15 years 0 months (from Start and Wells, 1972)

A The Watts-Vernon (WV) Test
(Maintained and Direct Grant Schools)

Date of survey	1948	1952	1956	1961	1971
Mean score	20·79	21·25	21·71	23·85*	23·46

*Estimated from the obtained mean applying to secondary modern and comprehensive school pupils only, being the mean of the *two* estimated quoted by Start and Wells.

B The National Survey Form (NS6) Test
(Maintained schools only)

Date of survey	1955	1960	1971
Mean score	42·38†	44·51	44·65

†Estimated from the obtained mean applying to England and Wales.

What the authors emphasize however are not the differences between the 1971 mean scores and those immediately preceding but differences from 'expected' or 'anticipated' mean scores resulting from extrapolation of earlier increases. Thus, for the 11-year-olds' performances on the WV test an 'anticipated' mean of 16·30 is forthcoming if the almost linear trend from 1948 to 1964 were continued to 1970 and it is the 'drop' from this extrapolated figure, i.e. 16·30-14·19 =2·11 points of score and equivalent to over 12 months of reading age, which is hailed as a cause for concern. (A similar 'drop' in the 11-year-olds' performance on the NS6 test is equivalent to over 5 months of reading age.) Such differences between obtained and 'anticipated' means lead the authors to 'face up to the distinct possibility that the 1970/71 scores are in fact depressed'.

Now despite the cautious phraseology—in the realm of statistical estimation nothing can be ruled out with complete certainty—and the qualifications subsequently made, this conclusion merits a firm rejection as blatant nonsense. The only valid conclusion is that the null hypothesis of no real change in mean attainment since the preceding survey has to be retained. In particular it is wrong to describe a position of no change as one of decrease merely because a forecasted increase has not occurred. And the description is especially unfortunate when there is no rational basis for the forecast. This last point is taken up again later.

There are moreover two further reasons why the pessimism of the authors is misplaced. Both are discussed in the report but are discounted as of little weight. One concerns the ageing of the tests. At the time of the investigation the WV test was 23 years old and the NS6 was 16 years old, and both were showing signs of being dated. The authors draw attention to certain items which would obviously not find a place in a newly constructed test, and point to the possibility of the pupils of the latest survey being at some disadvantage relative to those tested earlier. This possibility is discounted however since the 'ageing' was twice as long for the period 1948-64 (when the previous survey was undertaken) as for the period 1964-70, and that it would therefore be necessary to postulate 'extreme change' in the latter period. But the effect could very reasonably be expected to become more intense, for a given time span, the more distant the date of the test. An age increase of three years could have little effect on the footballing skill of a player in his middle 20s while the same increase in his early 30s could be catastrophic. Familiarity with language, and within a particular social climate, may well be subject to a threshold effect in much the same manner. Again the depression of test score induced by ageing could follow a logarithmic rather than a linear trend. But concrete evidence on this point would be welcome. An estimate of the effects of dating could be obtained by separating off the obviously out-dated items and comparing the change of score on the 'out-dated' and 'less out-dated' sections of the tests. Perhaps the authors could be persuaded to publish a research note on this?

A second reason why a pessimistic interpretation of the findings is to be deplored derives from the 'immigrant corrections'. The proportion of immigrants (almost all low scorers on the tests) had increased substantially since the previous surveys, and the authors estimate that if the West Indian, Indian and Pakistani children had not participated, the mean score of the junior group would have increased by 0·15 points on the WV test and by 0·18 points on the NS6 test. For the senior group the immigrant corrections would have been higher, approximately 0·3 on the WV test and 0·8 on the NS6. These corrections were not in fact applied as it is possible they are over-corrections, since *some* immigrant children had taken part in the previous surveys. But why could not corresponding corrections have been applied to these results also? A second reason advanced for not applying these corrections to the 1970/71 data is that the immigrants are now better taught, and so presumably are less deserving of treatment as a special sub-group. But this is a *non sequitur*. Comparisons, if they are to be meaningful, must be between like and like. It is worthy of note that if the above immigrant corrections had been applied, the drop in the junior WV mean (Table 3) would not have been significant even if corrections to the age midpoint had been applied.

Comparisons of the proportion of illiterates and semi-literates (based on the 1938 criteria reference levels for reading ages of seven and nine years) would need similar qualifications, and especially with respect to the vastly increased numbers of immigrants in the 1970-71 survey. Apart from this it needs to be pointed out that the standard errors of percentages derived from small and extreme subgroups are very considerably greater than for the overall group means, and would inevitably cloud into statistical insignificance the differences reported. The reporting of figures showing the change in such percentages without reference to the increased statistical errors involved can only be deplored.

(e) *Evaluation:* The high proportion of schools in the design sample refusing to co-operate renders suspect the application of the findings to the country as a whole. It is unfortunate, too, that an unjustified inference of decline in standards could well be made because of the incorrect basis of comparison adopted and other reasons.

The Cockburn Study (1973)

(a) *Aims:* The main aim is that of determining the extent of reading backwardness among seven-year-old children in Angus, Scotland, and of relating the findings to those of previous annual surveys. It was also intended to assess the reading attainments of a subgroup who had been identified as backward readers three years ago.

(b) *Criteria:* The *Vernon Graded Word Reading Tests* had been used for all annual surveys since 1963. This is an individually administered test of word recognition and has Scottish norms. For the assessments of those previously identified as backward readers the *National Survey Form Six* (NS6) test was used.

(c) *Sampling:* All seven-year-old children attending schools in Angus in each of the years 1963 to 1972 were tested. (Testing had been undertaken in 1962 but a different test was then used.) The testing was undertaken in all cases by the class teacher, so any absentees were presumably tested on their return to school. Some 1,600 children were involved in all in 1972.

In addition a survey was made in 1969 which included all the seven-year-olds tested in 1966, and who were then found to be backward readers. 117 children were in this category.

(d) *Results:* The percentage of seven-year-olds classified as backward readers by their reading age, as determined by the Vernon test, being at least one year below their chronological age, was found for each of the years 1963-72. The results are shown in Table 5. We see that the percentage of backward readers rises sharply in the years 1968 to 1970 and this is only partially offset in 1971 and 1972. Cockburn comments that the poor results of 1968-70 were not confined to a small number of schools, or to any one type of school, but were general, affecting town and rural schools alike. The difference between the percentages for 1968 and 1966 (selected for this purpose presumably because it is the lowest) was found to be significant at the one percent level. Apart from this being an *aposteriori* comparison, and therefore suspect when tested by the same procedures as for planned comparisons (Edwards, 1966, 329-330; Lewis, 1968, 49-52; Snedecor and Cochran, 1967, 268-271), it is debatable whether any differences between years should be tested for statistical significance at all. What, after all, is the 'population' with respect to which generalization is desired? This does not detract of course from the value of the data both intrinsically and as a basis for possible changes in educational practice.

In the survey of ten and 11-year-old children, only 88 of the 117 backward seven-year-olds had remained in the county and of these all but eight were still backward readers. They were joined in this respect by another 43 who had not been identified as backward readers at the age of seven.

(e) *Evaluation:* This is a valuable contribution to our knowledge of reading standards and is precisely the kind of exercise, involving the cooperation and participation of teachers, that needs to be repeated in other parts of the country. An increase in the proportion of backward readers in Angus since 1967 has been established. No information is provided as to whether this is due primarily to population changes in the county or to different conditions in the schools.

Concluding discussion

All the studies reviewed in this article are essentially normative in that they seek to ascertain standards, regional or national. At the same time some are also concerned (and possibly to an at least equal extent) with a probing for variables of an organizational or psychological kind associated with relatively high attainment (or, negatively, with a relative absence of low attainment) in reading. We see too that almost all employ either a single criterion or two criteria of the same type, and that even when two or more essentially different measures of reading are

Table 5: Percentage of backward readers in the Angus Annual Reading Surveys (from Cockburn, 1973)

Year of survey	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
Percentage	9·9	10·4	8·6	8·4	11·0	13·5	14·6	15·0	12·9	12·7

employed, these are treated separately, and not in combination. In other words all the studies are uni-criterion in design. Conclusions from even the most ambitious study—ambitious as far as geographical coverage or intensity of sampling is concerned—must therefore be modest. This is not to deny the relevance of the type of criteria chosen. The tests of reading comprehension and, in the case of younger children, word recognition might well impress as reasonable, and set against the total complex of reading skill could emerge with respectably moderate validities. At the same time the confinement to the cognitive is worthy of note and should perhaps be contemplated with at least a modicum of unease.

Turning now to sampling methods—the second basic issue mentioned in the introduction—we see that the studies divide into three groups of two. In the ILEA and Cockburn studies practically the whole of the defined population is sampled, with the need to generalize the resulting statistics consequently removed. In a second group, consisting of the studies of Kellmer Pringle *et al* and Clark, the child is the sampling unit, while in the third group of studies, those of Morris and Start and Wells, the initial unit of sampling is the school. Insufficient information is provided to be sure that for this last group account of the two stage sampling has been taken in extracting the estimates of statistical error.

Another fundamental difference among the studies is that of the geographical coverage attempted, and here the division is between the two national surveys on the one hand and the four regional (or county) surveys on the other. Yet another distinction is whether or not the findings are placed in relation to those of previous surveys, in a time sequence, so that a rise or decline in standards may be inferred. A common problem here, and one discussed with reference to the Start and Wells study—is that of the ageing of the tests. The best solution would seem to be for a test suspected of ageing to be used only alongside a new replacement test on the next survey, the replacement test then taking over (until it in turn becomes 'aged') on subsequent surveys. Resorting to subdivisions of the test, to 'aged' or 'dated' and 'less dated' sections as suggested in the criticism of the Start and Wells study, is obviously very much a second best.

The Start and Wells study can be placed in a category of its own if only because it is both a national survey and one which places its findings in a temporal sequence. Again it is the only study reviewed that deals with standards of reading in secondary schools. Attention has been drawn to its defects and in particular to the use of a forecasted increased score as a baseline for assessment. It is interesting to speculate why such a baseline was chosen. It may well be part of the same climate of thought that exalts economic growth—as assessed by a percentage increase in gross national product—above all else. In the present context such thinking would lead to an expectation of increased performance in tests of the comprehension of incomplete sentences even though the validity of such tests becomes increasingly imperfect for older children. It is at least arguable that the benefits of any increased 'input'—in the form of more enlightened teaching for instance—should be expected in this particular way. A more comprehensive approach to the basic problem of criteria together with an increased use of regional (or local authority) surveys would seem to provide the most promising lines of advance at present.

References

Barker Lunn, J. C. (1970). *Streaming in the Primary School*. Slough: NFER.

Brimer, M. A. and Dunn, L. M. (1963). *English Picture Vocabulary Tests*. Bristol: Educational Evaluation Enterprises.

Butler, N. R. and Bonham, D. G. (1963). *Perinatal Mortality*. Edinburgh: E. S. Livingstone Ltd.

Campbell, D. (1951). 'Facts relevant to the validity of experiments in social settings', *Psychol. Bull.*, **LIV**, 297–312.

Campbell, D. and Stanley, J. (1963). 'Experimental and quasi-experimental designs for research on teaching', Chapter V of *Handbook of Research on Teaching* (Gage, N., Editor). Chicago: Rand McNally & Co.

Clark, M. M. (1970). *Reading Difficulties in Schools*. Harmondsworth: Penguin Papers in Education.

Cockburn, J. M. (1973). 'Annual surveys of reading disability in a Scottish county', *Brit. J. Educ. Psychol.*, **43**, 2, 188–191.

Edwards, A. L. (1966). *Statistical Methods for the Behavioural Sciences*. New York: Holt, Rinehart and Winston.

Inner London Education Authority (1969). *Literacy Survey: Summary of Interim Results of the Study of Pupils' Reading Standards*. Cyclostyled Report from the Education Officer, The County Hall, London.

Kellmer Pringle, M. L., Butler, N. R. and Davie, R. (1966). *11,000 Seven-Year-Olds: First Report of the National Child Development Study (1958 Cohort)*. London: Longmans.

LEWIS, D. G. (1968). *Experimental Design in Education.* London: University of London Press.

LEWIS, D. G. (1972). *Research Methods in Education: The Basic Concepts of Design.* London: Thames Polytechnic Address.*

LINDQUIST, E. F. (1956). *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mufflin.

MCNEMAR, Q. (1962). *Psychological Statistics* (3rd edition). New York, London: John Wiley & Sons.

MORRIS, J. M. (1966). *Reading in the Primary School.* London: Newness.

MORRIS, J. M. (1966). *Standards and Progress in Reading.* Slough: NFER.

NEALE, M. D. (1966). *Neale Analysis of Reading Ability.* London: Macmillan.

SCHONELL, F. J. and SCHONELL, F. E. (1960). *Diagnostic and Attainment Testing.* Edinburgh: Oliver and Boyd.

SNEDECOR, G. W. and COCHRAN, W. G. (1967). *Statistical Methods.* Ames, Iowa: Iowa State University Press.

SOUTHGATE, V. (1962). *The Southgate Group Reading Tests: A Manual of Instructions.* London: University of London Press.

START, K. B. and WELLS, B. K. (1973). *The Trend of Reading Standards.* Slough: NFER.

*Copies available on request from the Department of Education, University of Manchester.

RESEARCH IN EDUCATION

May 1975 *Number 13*

The impact of social science on flows from school to university: D. HUTCHISON and G. M. LITTLEJOHN

Objective test analysis: some criteria for item selection: J. WILMUT

The satisfaction of university staff and their teaching: RICHARD STARTUP, MICHAEL GRUNEBERG and PATRICK TAPSFIELD

Syllabus-bound and syllabus-free orientations in a technological university: ALAN SMITHERS, PAT SAYLES and ALAN JOSEPHS

Some thoughts on norm-referenced and criterion-referenced measures: T. J. HARVEY

Published twice-yearly, May and November
£4.50 per annum, single copies £2.75

MANCHESTER UNIVERSITY PRESS
Oxford Road, Manchester M13 9PL

Measurement of Human Values:

an Appraisal of the Work of Milton Rokeach

Tom M. Kitwood and Alan G. Smithers, School of Research in Education, University of Bradford

Summary

It is argued that methods currently available for studying values within a socio-psychological framework are inadequate and that, if, for example, the interrelationships between values and education are to be systematically explored, then it is essential that approaches be devised which reflect more accurately the nature of human valuing. Some of the problems associated with present methods are illustrated by examining in detail one of the latest and, in some ways, most impressive to become available—the Rokeach Survey of Values. It is suggested that, in spite of its simplicity and ready appeal, it suffers from important weaknesses (notably its inadequate conception of human values) which are liable to seriously distort the information which it appears to give. Some of the requirements for an adequate questionnaire method of studying values are outlined, but the importance of obtaining corroborative evidence from idiographic and field approaches is also stressed. The paper illustrates the kind of basic thinking required before values are studied or measured.

Introduction

Although it is widely recognized that systematic knowledge about personal values is crucial to the understanding of human behaviour, methods for the study of values within a social-psychological framework are as yet relatively poorly developed. Issues concerning values are highly relevant to education, where there appear to be two central questions.

The first is: should educators deliberately attempt to inculcate values into those with whom they are dealing? This is a question for philosophers of education, and around it there has been controversy from the time of Plato to the present day. The second question is an empirical one: to what extent, in fact, are personal values affected by education, by the overt or the hidden curriculum? It seems generally accepted that the school leaver or graduate emerges with a value system of some kind, acquired in some process; and it is often assumed that he will keep it, unchanged except in minor ways, for the rest of his life. The extent to which education is involved in the formation of value systems still awaits thorough investigation. If that and other vital questions are to be answered, adequate methods for studying values are essential.

One of the problems here is that the methods which are available have embodied different conceptions of personal values. The earliest, and probably the best known, is the Allport-Vernon-Lindzey Study of Values (1960) for which Buros (1972) lists 687 references. Forms of this test have been used for over 30 years. However, since it has been designed to assess 'the dominant *interests* of the personality' (our italics), it really does not measure values as such, for while values pervade all interest areas, they are neither synonymous nor co-extensive with them. Another approach, Morris's (1956) Ways-to-Live study, which might have led to much further work, but which now seems largely forgotten, measures values as preferred life styles. The famous Value Orientations study of Kluckhohn and Strodtbeck (1961) treats values as answers to certain fundamental questions about the nature of reality. Gordon's Survey of Interpersonal Values (1960) and Survey of Personal Values (1964) overlap considerably with the concept of 'need' adopted by workers such as Murray (1938). These and a number of other studies are reviewed in Kitwood (1974).

Milton Rokeach, probably more than any

other living psychologist, has appreciated the centrality of values in the understanding of human behaviour. He has been concerned to avoid the superficialities present in similar areas of social psychology, notably some of the work on attitude and attitude change. His approach to the value question is very direct, and is based on a conception of value which is, arguably, more accurate than those of earlier workers.

In *The Nature of Human Values* (Rokeach, 1973), which brings together much of his work to date, he defines a value as 'an enduring belief that a specific mode of conduct or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct or end-state of existence. A *value system* (Rokeach's italics) is an enduring organization of beliefs concerning preferable modes of conduct or end-states of existence along a continuum of relative importance'. He thus follows the familiar ends-means dichotomy and divides values into those concerned with ends, 'terminal values', and those concerned with means, 'instrumental values'.

The Survey of Values

These conceptions are operationalized in the Survey of Values (Rokeach, 1967) in which the respondent is simply invited to rank order 18 'terminal values', concerned with end-states of existence, and 18 'instrumental values', concerned with modes of conduct. Each 'value' is presented as a word or brief phrase, with a few words of explanation. Typical examples are: Family Security (taking care of loved ones), Pleasure (an enjoyable, leisurely life), Capable (competent, effective), Obedient (dutiful, respectful). Of these, the first two are 'terminal' and the latter two 'instrumental'. One of the novel features of the Value Survey is that the items are printed on small removable gummed labels. Thus the respondent is free to adjust them until he has produced an order which he believes represents his value system. This is a very original way of presenting test material. No writing is involved; the task appears to be interesting; the time to complete it is short, about 10-20 minutes; the completed Survey is, in effect, just a series of numbers, a body of information very easy to transfer and record.

The Value Survey has already proved a versatile instrument, having been extensively used by Rokeach and his co-workers, and by Feather (1972) in Australia, and it may well come to form the basis of a new orthodoxy in value theory. However, despite the thoroughness with which it has been developed, and its simplicity and ready appeal, we believe that it suffers from several important weaknesses. A discussion of them illustrates some of the problems which are involved in the study of human values, an area in which distortion can very easily occur.

1. From the outset Rokeach assumes that there is a strict rank-ordering of the elements within the personal value system, and this underlies his whole approach. There is, however, no theoretical reason for supposing this to be the case. While it is likely that some values are more important than others for any one person, this does not imply that they must form a strict hierarchy. The only condition for values to be rank-ordered is if the holding of one automatically excluded the holding of another at the same level. This is the case when two values imply alternative answers to the same questions, and hence alternative courses of action. Thus it is possible to conceive of a decision related to a choice between the values of An Exciting Life and Family Security, but it is very hard to visualize a situation where the values of Inner Harmony and Wisdom could be set in mutual opposition. The problem here is made more serious by the fact that the 36 'values' are not all of the same type. Rokeach himself makes the point that some of his 'terminal values' are 'personal', some 'social', and suggests that the 'instrumental' list involves both 'morality' and 'competence'. He does not seem to have reckoned sufficiently with the fact that his values imply answers to questions of several kinds, such as 'What kind of person do I want to be?', and 'What kind of world do I want to live in?'; the answers to the different questions do not lie on the same continuum. Thus the invitation to rank Self-Respect, A World at Peace, and True Friendship is about as meaningful as to ask, 'which do you prefer, strawberries, Bach, or air travel?' Although no doubt many people readily carry out the ranking task as offered, this does not justify a basic conceptual weakness. The fact that the respondent is willing to 'play the game' is no guarantee that the game is logically valid.

2. The Value Survey, like many of its predecessors, is ipsative; that is, the respondent is

asked to arrange a finite number of units within an enclosed framework, provided by the questionnaire itself. The task is limited to making adjustments within that framework. While ipsative methods can be justified in some areas of psychometrics, they seem to involve unwarranted assumptions in the measurement of values. To use an ipsative method is to assume that all people have the same structure and dynamics in their value system. Thus Rokeach's method implies that all people have the same endowment of 'value energy' to be spread over 18 instrumental and 18 terminal values. An ipsative method such as this provides no way of investigating interpersonal variation in what may be very significant aspects. It is possible that a person with very little 'value energy' might complete the Value Survey identically to a person who held his values with great intensity; that one whose value system was crude and undifferentiated would be indistinguishable from one whose value system was subtle and coherent.

3. There are also problems with the items themselves. These are presented as very brief phrases, open to a number of constructions by the respondent. The test thus has a considerable projective element in it, but without any means of examining what projective processes have taken place. The 'value' of Freedom, or of Mature Love is treated as if it meant the same to all respondents, whereas Rokeach may in fact be measuring some kind of affective response to the aura and associations of the verbal stimuli. Rokeach claims that the Survey can be used with children as young as 11 and has reported a number of patterns for the development of values; but for how many young people would Social Recognition or National Security have a clear meaning? The ambiguity of the values also throws doubt on Rokeach's work on the modification of values. He has reported that, by bringing into the open apparent inconsistencies between a person's self-conceptions and values, it was possible to induce changes in certain 'target' values which persisted for 15 months or more. However, it is possible that during the experiments the subjects simply became more clear about what Rokeach himself meant by the value items.

4. The actual choice of values presents some surprising anomalies. Among the terminal values there is no mention of Truth, which must be reckoned to be one of the highest for those committed to science, philosophy, the arts, and religion; another omission is anything to do with health, vitality, and the full use of personal powers. Conversely, considering the limitation of numbers, it is surprising that Comfort and Pleasure, which empirically were found to have the highest mutual correlation between any two of the 36 values, are both included.

Another cause for questioning here is the instrumental/terminal dichotomy, at least as operationalized by Rokeach. The distinction between Happiness and Cheerful, between Mature Love and Loving, between Freedom and Independent (the first being 'terminal' and the second 'instrumental' in each case), is scarcely tenable. This is borne out by a factor analysis reported by Rokeach in which values of both kinds were found on each factor. The first factor, for example, 'immediate *v* delayed gratification' had positive loadings from A Comfortable Life, Pleasure, Clean, and An Exciting Life, and negative loadings from Wisdom, Inner Harmony, Logical, and Self-Controlled.

5. The ranking task appears conceptually difficult. Miller (1956) has discussed the number of categories which the human mind can hold simultaneously, suggesting that this is seven, plus-or-minus two. Rokeach acknowledges this, and accepts a limit of about nine items in a strict ranking task. He comments thus about the task in his Value Survey: 'The respondent's task becomes progressively lighter with every gummed label that he ranks . . . after the respondent had moved twelve values . . . he has only six more to compare with one another . . . thus the average length of the scale turns out to be 9·5'. This is not a very convincing argument; the task is, in fact, a re-ordering of a scale 18 items long, and it seems fairer to say that in terms of Miller's limit about half the task is within a person's conceptual powers and about a half beyond them. This, together with the difficulties concerning the nature of the items, may account for the rather low reliability coefficients (average around 0·65) reported by Rokeach.

6. Finally, a question must be raised about the way in which values are related to beliefs, attitudes, and actions. According to Rokeach, whereas a person may hold tens of thousands of

beliefs, and thousands of attitudes, the number of values is quite small. A value 'has a transcendental quality to it, guiding actions, attitudes, judgements, and comparisons across specific objects and situations and beyond immediate goals to more ultimate goals. They are the determinants of attitudes as well as of behaviour'. Statements such as these incorporate a number of assumptions, which at least deserve investigation. In the case of the first, there may well be a confusion between logic and psychological mechanism. While it is easy to demonstrate that generalized value propositions are logically prior to attitudinal propositions, this does not imply psychological causation. The function of a value proposition in logic is not necessarily related to the function of values for the person. Similarly, while Rokeach speaks of values as 'determinants of behaviour', and has investigated this question at a gross level, there is as yet little evidence about the way values are related to the microstructure of human action, to individual decisions and their implementation. Here questions of relevance and competence are also involved.

Possible future Directions

If values are to find a satisfactory place in a social-psychological framework, at least the following should be taken into account: the presence of differing value conceptions, both as concerning obligation and satisfaction; variations in capacity to handle values as concepts, and, in particular, to assign priority; variations in ability to relate values logically to attitudes and decisions; variations in the stability of value systems over a range of social situations. Thus the difference in 'valuing' between persons may be very wide, and not adequately accommodated within any of the existing instruments.

In order to achieve an approximate correspondence with reality, it seems to us that a questionnaire method of measuring values would have to meet the following requirements:

(a) It should be based on a clear definition of the value area, in which both obligation and satisfaction (moral and utilitarian values), are given due weight. The value concepts should be as free as possible from class and culture bias.

(b) It should be conceptually simple, so that it can be used in the lower intelligence range as well as among those whose intellectual powers are high. This means that the best kind of item is probably that which involves 'realistic decision choice' i.e. making reference to concrete situations, rather than abstractions.

(c) It should be non-ipsative, so that it is capable of measuring variation in intensity, and in particular be sensitive to low levels of valuing.

(d) The constraint on the respondent should not be too high. In expressing his personal values the respondent should be given as much freedom as possible.

Even if an instrument worked out on a theoretical foundation and meeting these requirements could be devised (and since the requirements are not mutually incompatible there is every reason to suppose that it could), results obtained by it would have to be treated with considerable caution, and corroborated by other approaches. There is a need for the development of an idiographic approach to value study, which makes minimal assumptions about the value elements and the ways in which they are organized and used by the person. Also, it is desirable to develop field approaches which would examine values in 'real-life' situations. However, here it would not be sufficient simply to quantify units of behaviour, since there are strong grounds for thinking that behaviour is rarely a direct expression of values (Kitwood, 1974).

Conclusion

So, because of its importance both to education and to the wider field of social inquiry, we recommend that those who try to study values do so with caution. There must be an awareness of what any instrument may fail to reveal because of the limitations incorporated into its design; also of what a set of apparent 'results' may actually obscure. It is essential that new approaches be devised, which have a more authentic bearing on the realities of human valuing. Some of these are being currently explored in our current work at Bradford.

References

ALLPORT, G. W., VERNON, P. E. and LINDZEY, G. (1960). *A Study of Values* (3rd ed). Boston: Houghton Mifflin.

BUROS, O. K. (1972). *The Seventh Mental Measurements Yearbook*. New Jersey: The Gryphon Press.

FEATHER, N. T. (1972). 'Value similarity and school adjustment', *Aust. J. Psychol.*, **24**, 193-208.

GORDON, L. V. (1960). *Survey of Interpersonal Values*. Chicago: Science Research Associates.

GORDON, L. V. (1964). *Survey of Personal Values*. Chicago: Science Research Associates.

Kitwood, T. M. (1974). *The Nature of Human Valuing*. Unpublished M.Sc. dissertation, University of Bradford.

Kluckhohn, F. R. and Strodtbeck, F. L. (1961). *Variations in Value Orientations*. Illinois: Row, Peterson and Co.

Miller, G. A. (1956). 'The magical number seven, plus-or-minus two: some limits on our capacity for process-information', *Psychological Review*, **63**, 81-97.

Morris, C. (1956). *Varieties of Human Value*. Chicago: The University Press.

Murray, H. A. (1938). *Explorations in Personality*. Oxford: The University Press.

Rokeach, M. (1967). *Value Survey*. Sunnyvale, California: Halgren Tests.

Rokeach, M. (1973). *The Nature of Human Values*. New York: The Free Press.

Journal of Curriculum Studies **Collins**

In the November issue

Secondary School Curriculum and Social Change in an Emergent Nation. *M. K. Bacchus*

Gagné's Hierarchical Theory of Learning: Some Conceptual Difficulties. *David Soulsby*

A Systematic Approach to Curriculum Evaluation. *S. A. Lutterodt*

A Learning Hierarchies Approach to Social Studies Instruction. *B. S. Deming*

A Comparison of Regular and Open Classroom Process. *D. W. Cochran and B. W. Tuckman*

The Teachers' Curriculum. *I. Goodson*

It may be ordered direct from the Publishers

In Britain
The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow

In the U.S.A. and Canada
APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

May and November
Annual Subscription £3.00 including postage

Can Pre-School Children Form Concepts?

Robert Povey and Eric Hill, Christ Church College, Canterbury

Summary

Fifty-six children between the ages of two years four months and four years ten months were given tests relating to the acquisition of both 'specific' and 'generic' concepts. Several tests of concept acquisition (HAPCAT)[1] were devised by the authors utilizing pictures drawn on card. A number of the children (N=29) were also given some 'Piagetian questions' concerning class inclusion.

The results conflict with the widely held view (following the writing of Jean Piaget) that pre-school children cannot form generic concepts. Nearly all the children were able to identify the specific concepts as presented in the HAPCAT items and about half the group responded appropriately to the HAPCAT items involving an understanding of generic concepts. There was a clear and statistically significant relationship between the number of correct answers given to the HAPCAT items and the age and ability levels of the children. This was not the case, however, with the Piagetian test questions which showed no discriminatory power in these respects. It is argued that the Piagetian questions do not represent a sensitive or meaningful measure of concept acquisition at this age level whereas the HAPCAT items do appear to present pre-school children with intelligible tasks which allow many of them to demonstrate an understanding of generic concepts involving class inclusion.

Background

According to various Piagetian interpreters (e.g. Beard, 1969; Stones, 1966) children are not capable of forming 'true concepts' during the pre-school stage. Describing the 'pre-conceptual' phase of development (two to four years) Stones, for example, states that this is the phase where 'the child is beginning to develop concepts and before he is capable of distinguishing between the specific concepts relating to an individual object and the generic concept relating to a class of objects'. He gives as a common example of this 'the tendency of children of this age to call men generally "daddy"'. Similarly, Beard claims that children's verbal concepts in the pre-conceptual phase 'lack the generality of true concepts'.

At a point even further removed from Piaget's own writing one finds students' interpretations of the interpreters. Witness, for example, the following comments which we have read in the answers of quite able students to questions on Piagetian researches (the comments having been selected from papers by BEd and Teachers' Certificate students from different colleges):

> 'Concepts can only be acquired from four onwards.'
>
> 'Pre-conceptual (two to four years) where everything on four legs is a dog and any man is "daddy".'
>
> 'Not until the age of four or five will the child distinguish between the characteristics which mark "dog" and those which mark other animals.'

Acknowledgements
The authors would like to acknowledge the most generous help and co-operation they received from Mrs Robin Dowie, Mrs Barbara Best and the mothers and children at the St. Stephen's Community Centre Play Group, in Canterbury.

---

[1] The tests may be obtained from the NFER (Guidance and Assessment Service) to whom inquiries should be addressed.

> 'No distinction between generic and specific terms. Little point in teaching other than play.'

We believe that these extracts from the writing of 'authorities' and students are fairly typical examples of a widely held belief that children at the pre-school level are not capable of forming 'true concepts' and, in particular, that they cannot form generic concepts.

Piaget (1962) elaborates this belief in relation to the development of concrete operations. He states that hierarchical classification involving logical inclusion is not possible before about seven or eight years of age: 'the simplest operation is concerned with classifying objects according to their similarity and difference. This is accomplished by including sub-classes within larger and more general classes, a process which implies logical inclusion. Such a classification, which seems very simple at first is not acquired until around seven or eight years of age'. Other writers, however, such as Russell (1956) claim that normally children have been discriminating, abstracting and generalizing about environmental data from infancy and that by the age of three or four a child 'knows literally hundreds of concepts'. Welch (1940) also claims quite explicitly that pre-school children are capable of forming hierarchical concepts involving class inclusion.

Such apparently diverse views may be reconciled partly by reference to differences in the interpretation of the term 'concept'. There is a certain amount of confusion in the literature over the use of the term 'concept' in the sphere of intellectual development (see Blank, 1968) but one can also find a considerable degree of general agreement that the acquisition of a concept in its more usual connotation is concerned with abstracting the essential attributes of an object or idea, distinguishing this from other objects or ideas and generalizing to new situations. Brown (1967) put it this way: a concept is 'a way of grouping an array of objects or events in terms of those characteristics that distinguish this array from other objects or events in the universe . . . and a concept is considered to have been attained when a subject is able to identify new instances of it without further training'. If one accepts this meaning of the term 'concept' then the development of concepts is not, of course, concerned only with hierarchical classification. Thus statements which claim that children are capable of forming 'hundreds of concepts' by the age of three of rour (Russell, 1956) may be partly reconcilable with the Piagetian viewpoint on the grounds that they are referring to concepts at a different hierarchical level.

Such an explanation, however, does not meet the claims put forward by Welch (1940) concerning the acquisition of *generic* concepts by pre-school children. He describes several hierarchical levels of concept acquisition. The first he calls the 'pre-abstract' stage at which the child learns the attributes of one class of objects (e.g. 'apples')—the concepts at this stage are specific and not generic. The child then learns that one class may include another e.g. 'fruit/apple' (the first hierarchy) and, in turn, that this class may include other classes e.g. 'food/fruit/apple' (the second hierarchy) and so on. Welch claimed that children are able to form concedts at the 'pre-abstract' level (e.g. 'chair', 'ball') before the age of two years and without the child using speech. The tests for the acquisition of such concepts involved the child responding appropriately to the test words by pointing out a chair or ball from among a group of other objects, and then generalizing to new situations (e.g. the presence of a different shaped chair or a folded chair). Similarly, he suggests that by the age of about 26 months the children were able to understand some concepts involving the existence of a sub-class within a more general class (i.e. first hierarchy concepts). Thus, according to Welch, many children at this age understood that 'men' and 'women' are also 'people' and that 'potatoes' and 'apples' are both 'food'. Further, he suggested that children tend to acquire *second* hierarchy concepts rapidly at about four-and-a-half years of age.

There is a convincing relationship between many of Welch's conclusions and the straightforward observation of children's behaviour. Most parents would probably attest that children between the ages of two and four tend to have a fairly secure grasp of concepts such as 'cup', 'spoon', 'cot', 'dog', etc. These are specific concepts certainly and do not involve logical inclusion but they are nevertheless 'true concepts' on the definition which we have discussed above i.e. the child has grasped the essential attributes of the object or idea, can distinguish this object or idea from others and can generalize to new situations. Similarly, parents and nursery teachers may well wonder at Piaget's assertion that children in the pre-school stage are not admitted to an understanding of some generic concepts involving class inclusion, when they find their children using generic terms such as 'clothes' and 'food' in an appropriate manner and responding appropriately to the adult's use of such terms. Thus there seem reasonable grounds for hypothesizing that many children in the age group usually taken to coincide with the pre-conceptual stage (*viz.* two to four years) are able to form generic concepts. For example, one might suggest with some confidence that pre-school children might be able to recognize and distinguish between a number of different kinds of food (e.g. ice-cream, apples, potatoes) and at the same time understand that these also constitute a further class of 'food'; and if they do have the capacity for such understanding then they have grasped something about hierarchical classification involving class inclusion.

Despite such considerations, Piaget and Inhelder (1964) dismiss the view that children are able to form hierarchical concepts at the pre-conceptual stage. They claim that Welch's work merely demonstrates that children in this phase are capable of forming 'non-graphic' collections although they do not appear to dispute Welch's argument that pre-school children can form concepts at the 'pre-abstract' level. In Piaget's view the child progresses towards full hierarchical classification in the following stages: from '*graphic collections*' where children place the objects together because they seem to 'belong' e.g. children might group a baby with a cot, instead of classifying the baby with people and the cot with furniture; through the second stage of '*non-graphic collections*' in which objects are assigned to one collection or another on the basis of similarity alone without any notion of hierarchical classification involving class inclusion; and finally to classification involving inclusion.

According to Piaget the understanding of class inclusion can be assessed most precisely by the use of certain questions, especially those involving the terms 'all' and 'some'. The 'all/some' questions are designed to elicit comments on the relationship between a sub-class ('some') and an enveloping class ('all'). A further set of questions relate to a situation in which one class A (e.g. primroses) is included in another class B (flowers) without being equal to the whole of B—the question is then asked, 'Are there more As than Bs or more Bs than As?'. Although we have not set out specifically to examine this point, we would doubt the strength of Piaget's claim that these types of question are, in fact, adequate tests of a child's understanding of hierarchical classification.

It might be argued (and Piaget and Inhelder, 1964, do in fact discuss this point briefly) that an adequate understanding of these questions relates rather more closely to verbal facility and comprehension than to an understanding of hierarchical classification. Consider, for example, the question 'Are all the primroses some of the flowers?' For subjects of any age this is a distinctly odd question! Certainly in English it does not represent a linguistic structure with which children would be at all familiar and it does not seem entirely surprising that children find some difficulty in knowing how to answer. It is interesting that in Piaget's own experiments with children from six to eight years of age 'only 21 per cent of the subjects accepted statement (A) 'All the tulips are some of the flowers', while eighty-one per cent accepted the logically identical statement (B) 'Some of the flowers are tulips' (Piaget and Inhelder, 1964). One interpretation of this result on any commonsense basis would seem to be that the problem does not really make an ordinary sensible question when phrased as statement (A) whereas the second phrasing, statement (B), fits in quite sensibly to normal linguistic structures and is accepted and dealt with in a straightforward way. Piaget appears, however, to ignore such a straightforward explanation of the findings.

Similarly, if you give a child a bouquet having six daisies and six other flowers and then ask, 'Are there more daisies or more flowers?' this question again must appear rather surprising. It savours very much of the trick question and seems nearer to certain tests of reasoning ability than of concept acquisition. Moreover, the distinction drawn in the question between 'daisies' and 'flowers' tends to suggest that the questioner has created an artificial distinction between the daisies and the other flowers for the purpose of the question and may appear to indicate that he wishes the child to answer in the same vein.

It seems possible, therefore, that the type of tests Piaget has been using as the crucial criteria in determining the stage at which hierarchical classification emerges in children may in fact be less than adequate and may give a misleading impression of the stage at which children are capable of handling hierarchical classification. We incline to the view that many children in the pre-conceptual age range (two to four years) when tested by the more traditional tests of concept acquisition (such as those utilized by Welch, 1940, and described earlier) may show a reasonably certain grasp of class inclusion, as suggested by Welch. By 'traditional tests of concept acquisition' we refer to those tests which attempt to assess the degree to which the child has shown the capacity to abstract the essential attributes of a given generic class, to distinguish between this class of objects or ideas and other similar classes and to generalize to a new situation. Where a child demonstrates this capacity, it seems reasonable to argue that the concept has been acquired as adjudged by fairly rigorous and meaningful standards. It is still possible, of course, that some children who had attained this level of achievement might be unable to retain the stability of the concept in certain conditions e.g. when subjected to the Piagetian line of test questions. It would suggest, however, that the standard textbook statement that children between the ages of two to four years are incapable of forming 'true concepts' is immensely misleading. Such statements may also have unfortunate consequences when one considers their effect upon the outlook which students (especially trainee teachers) may adopt in relation to their pupils (e.g. the possibility of seriously underestimating the capabilities of children at the nursery and infant school stages).

The present investigation was initiated in order to test our contention (supported by the work of Welch, 1940) that pre-school children can form both specific *and* generic concepts. We used as a basis for this study the definition of a 'concept' given by Brown (1967) and set out earlier in this paper. We also tested for concept acquisition by more traditional methods than those adopted by Piaget and his co-workers (Piaget and Inhelder, 1964) although we did use the series of Piagetian test questions in part of the investigation for comparison purposes. The study examines two main questions in relation to our sample comprising children of two, three and four years of age (i.e. at the age level usually taken to coincide with the pre-conceptual stage in intellectual development):

(1) To what extent had the children acquired certain specific (i.e. non-generic) concepts?

(2) Utilizing the specific concepts referred to in question (1), to what extent were these pre-school children capable of showing an understanding of generic or class concepts?

Experimental Procedure

The subjects

Fifty-six British children in the 'pre-conceptual' age range (*viz.* two to four years). There were 32 boys and 24 girls. The youngest child tested was two years four months and the oldest four years ten months; the mean chronological age was three years nine months (SD=eight months). The children were attending a play group which drew upon a catchment area having a good representation from the different socioeconomic groupings. For example, at one extreme, there were children of unskilled manual workers living on a nearby council estate, whilst, at the other extreme, were children of university tutors. Using the criterion of 'father's occupation' as an indicator of the socioeconomic level of the home we categorized the sample into three broad groupings. Group I we termed 'Professional/Managerial', Group II 'Clerical/Skilled Manual', and Group III 'Semi and Unskilled Manual'. Fourteen children (i.e. 25 percent of the sample) were included in Group I, 28 (50 percent) in Group II and 14 (25 percent) in Group III. Again, the range of intellectual abilities seemed fairly representative of the normal distribution pattern. On the Peabody Picture Vocabulary Test, for example, the mean score was 103·9 with a SD of 16·99 points and a range of 66 to 141. The records of four children from the original 60 tested were excluded from the analysis of the results because of incomplete records and two were excluded because their parents were not British.

The tests

All the children were given two sets of test items, the Peabody Picture Vocabulary Test (PPVT) and the Hill and Povey Concept Acquisition Tests (HAPCAT). In addition 29 children were given some of the Piagetian test questions (Piaget and Inhelder, 1964).

The PPVT is a picture vocabulary test standardized in the USA giving norms for children from the age of two years three months to 18 years five months (Dunn, 1965). It has been found to be reasonably appropriate for use with British children (Moss and Edwards, 1960) and offers the advantage of providing a swift but at the same time fairly useful guideline to the level of verbal intelligence of very young children.

The Hill and Povey Concept Acquisition Tests (HAPCAT) were designed to assess the degree to which a child had acquired a certain concept—where 'level of concept acquisition' simply refers to the child's grasp or level of understanding of the concept. To assess the child's grasp of *specific* concepts he was asked to identify pictures of various objects (e.g. 'egg', 'girl', 'spoon'); and to assess the degree to which a child had acquired a *generic* concept he was asked to select the pictures which depicted a certain concept (e.g. 'food') from amongst a number of 'distractors'. As a further test of the child's level of understanding of the generic concept he was asked to say whether each of the various pictures was an example of the concept or not.

The 'generic tests' concerned 'food' and 'people'. Four pictures of specific items of 'food' were drawn in ink, each on a piece of 3¼in. square white card. Most of the pictures contained some colour. The four pictures of items of food were: egg, bread, cake, apple. Similarly, four pictures of 'people' were drawn. These were of a man, lady, boy and girl. In addition to these pictures four 'distractors' were drawn in relation to each of the generic concepts tested. For 'food' the following items were chosen as distractors: a spoon and a cup, as examples of objects a child uses when eating or drinking and might be expected to group with food if the child is only producing a 'graphic collection'; a flower—something which grows out of the ground and can be picked but, unlike an apple, is not to be eaten; a crayon, which again looks to be something which might be eaten. As 'distractors' for 'people' the following items were selected: a teddy-bear, an inanimate object which often has the status of a 'person' within the child's family; and three examples of animate objects which a child might confuse with people—a cat, dog and bird.

The 16 pictures were presented to the child on a sheet of cardboard (13in. by 13in.) divided into 16 squares (3¼in. by 3¼in.). The pictures were displayed in a randomly determined sequence as shown below:

Egg	Girl	Spoon	Bird
Teddy-Bear	Man	Flower	Bread
Lady	Cake	Cat	Crayon
Apple	Boy	Cup	Dog

The children were tested individually by one of the two authors. First of all they were told that 'we're going to play a game with some pictures' and asked to identify each of the pictures in turn. Correct identification of these original drawings of objects was taken as evidence that the child had acquired certain specific concepts. We argued that this afforded reasonable evidence of concept acquisition since the child had generalized his knowledge of specific concepts to this novel situation and responded appropriately by identifying the particular object when asked to do so, and had successfully distinguished between one specific object e.g. a cat, and similar but different objects e.g. a dog and a bird. We allowed a limited amount of latitude in scoring this test, giving credit for children who gave close approximations to the correct identification e.g. we allowed 'pencil' for 'crayon'. If the child was not able to identify a particular picture (usually because he mistook the actual drawing for something else) then he was told what it was meant to depict so that he would not be handicapped in his approach to the 'generic' test which followed.

The second part of the test concerned the acquisition of generic concepts. The tester said, 'Now I want you to give me all the pictures showing food.' The choices given by the child were noted and then all the pictures were replaced in their original positions. The tester then told the child 'I am going to ask you some questions about each picture. Let's look at each one in turn'. Starting with the left hand side of the top row the tester said, 'Now this is a picture of an *egg*. Are eggs food?' The same procedure was repeated for each picture in sequence from left to right along each row. Finally, the child was asked if he could tell the tester 'What food is?' (i.e. the child was asked for a definition). The same procedure was repeated for 'people'. If the child was able to pick out the instances of the generic concept e.g. 'food' from the distractors then we argued that this provides evidence that the child has attained a reasonably certain grasp of the concept involved. If the child was also able to answer the follow-up 'positive/negative identification' questions concerning each picture then this would give further support to the contention that the child had achieved a fairly certain grasp of the concept. The additional ability to offer a verbal definition of the concept would offer further support at an even higher level of cognitive functioning (e.g. vocabulary definitions are not introduced until year six on the Stanford-Binet) that the child had, in fact, grasped the concept.

Just over half the sample (N=29, age range 2-4 years) were also given some of the Piagetian questions designed to test the child's understanding of class inclusion (Piaget and Inhelder, 1964). These were the 'all/some' and the 'more than' questions described earlier. The questions were asked in relation to six pictures of people. The pictures were placed on the table as shown below:

1 Girl 3 Man 5 Boy
2 Lady 4 Lady 6 Lady

and the subject was then asked to identify the pictures in the order indicated (one to six). This procedure was intended to assist the child in establishing that there were equivalent numbers of 'ladies' and 'non-ladies'. In this way we hoped to avoid the introduction of a numerical bias in favour of either 'ladies' or 'non-ladies' which might have led to some confusion in the interpretation of the results, particularly to Question Seven ('Are there more ladies or more people?'). For example, if we had used in our presentation two pictures of 'ladies' and four 'non-ladies' then we could not have been at all sure that the child who answered 'more people' to Question Seven did, in fact, mean the fully inclusive group of 'people' (i.e. 'ladies' plus 'non-ladies'). He might have been referring simply to the group of 'non-ladies' which would have been numerically larger than the group of 'ladies'. With equivalent numbers of 'ladies' and 'non-ladies' we could be more certain that an answer of 'more people' was meant to indicate the inclusive group of 'ladies' plus 'non-ladies'.

The pictures were then placed in a line in the order—Man, Lady, Boy, Lady, Girl, Lady—and the child was asked the following seven questions: (1) Are all these people? (2) Are all the ladies people? (3) If you take all the ladies away, will there be any people left? (4) Are all the ladies some of the people? (5) Are all the people ladies, or are some of the people ladies? (6) If you take away all the people, will there be any ladies left? (7) Are there more ladies, or more people? With questions (3), (6) and (7) children who gave wrong answers were offered a practical demonstration with the pictures and then the question was repeated.

Test procedure

Every child took part in two testing sessions. On the first occasion the children were given the HAPCAT items and on the second, the PPVT. These test sessions were spaced one week apart. All the testing was carried out on an individual basis by the two authors during the normal play group sessions. In addition to this basic test programme we were able to test 29 children on the Piagetian questions four weeks after the first test session. After the same interval (i.e. four weeks after the first tests) 17 children were also given the HAPCAT items again in order to offer some indication of the reliability of the tests on a test/re-test basis. A few of the 're-test children' were also selected to take part in a very small-scale and rather 'raw' experiment. This consisted of an attempt to examine whether teaching these children the 'generic' concepts during an interval of four weeks from the pre- to post-test produced any discernible effect on the children's test performance.

Results

PPVT scores

On the PPVT the distribution of scores showed a fairly close resemblance to the pattern of standardized scores given in the test manual, apart from the rather greater proportion of children in the above average range (see Table 1).

Table 1: The proportion of children within each scoring range (a) in the present sample, and (b) in the standardization results

PPVT Standardized Scores	*N*	*Present Sample* %	*Standardization* %
125+	7	10·7	5
110–124	16	30·4	20
90–109	20	35·7	50
75–89	11	19·6	20
Below 75	2	3·6	5

HAPCAT scores

The results on the HAPCAT items will be presented in relation to the specific concepts first and then in relation to the generic concepts.

The *specific concepts* depicted by each of the 16 pictures were adjudged to have been acquired if the children were able to identify the pictures and respond to them appropriately by giving the name of the object. It was possible, therefore, for each child to score a maximum of 16 points on this test. Out of 56 children, 28, or 50 per cent, scored 16 points, i.e. they identified each picture correctly; 40 (or 71 per cent) were correct or made only one error and 47 (or 84 per cent) were correct or made one or two errors. There was a clear increase in capacity for correct identification with age although even with children below three-and-a-half years of age a high percentage of them (61 per cent) gained scores of 15 or 16. No child scored below ten points. It must be pointed out also that many of the errors appeared to be simply errors of recognition of the *drawings* rather than the objects themselves. We had, of course, taken measures to ensure that children of the age range being tested would recognize the pictures fairly readily, but some children still found other quite reasonable interpretations of some of the drawings (e.g. one child thought the 'crayon' was a 'rocket').

As far as the *generic concepts* of 'food' and 'people' are concerned, about half the sample were able to select the pictures of food and people correctly from the group of 16 pictures, rather more children (59 percent) choosing 'people' without error than 'food' (43 percent), (see Table 2). On the other hand, the children found it very much more difficult to answer the follow-up questions 'Are eggs food?', 'Are girls food?' etc. On this test only about one-third of the sample were able to give the correct answers, 30 per cent correct for 'people' and 36 per cent correct for 'food'. Taking correct performance on *both* the 'choice of pictures' test and the 'positive/negative identification' questions, then 30 per-cent of the sample achieved full scores on 'people' and 25 on 'food'. Three to four times as many children over the age of four gained correct answers on both tests as did children under four. Comparing the ages of children giving correct responses on both tests for at least *one* of the concepts one finds a highly significant relationship between age (four years and over as compared with under four years of age) and accurate test performance ($P<0{\cdot}001$ on a χ^2 2×2 test of association).

There is also a significant though less pronounced association between performance on these tests and verbal ability as measured by the PPVT (above mean scores/below mean scores). The association was significant at the $P<0{\cdot}01$ level on the χ^2 test. There was no significant association between performance on the HAPCAT items and the socioeconomic background of the children. The PPVT scores showed a slight relationship with socioeconomic background (children from socioeconomic groups I and II gaining a greater proportion of above mean scores on the PPVT than children from group III) but the association was not statistically significant.

Some interesting tendencies are evident from

Table 2: The proportion of children in each age range giving correct responses in the two HAPCAT tests

		PEOPLE				FOOD			
		Correct Choice of pictures		*Correct Pos./Neg. Ident.*		*Correct Choice of pictures*		*Correct Pos./Neg. Ident.*	
Age group	Total N	N	%	N	%	N	%	N	%
Under 3 years	7	2	29	1	14	0	0	0	0
3–00 to 3–05	16	6	38	3	19	5	31	3	19
3–06 to 3–11	9	5	56	1	11	5	56	3	33
4–00 to 4–05	13	10	77	5	38	7	54	6	46
4–06 to 4–11	11	10	91	7	64	7	64	8	73
Total	56	33	59	17	30	24	43	20	36

the children's individual records. For example, over half the errors made by children in their selection of items representing 'people' concerned the inclusion of the 'teddy-bear' or the pictures of animals or both in the list of 'people'. Similarly, one-third of the errors in the selection of items of 'food' related to the inclusion of the article used for eating or drinking (*viz.* the pictures of 'cup' and 'spoon'). In Piagetian terms the children are forming 'graphic collections'. Another feature of some children's performance was their tendency to start off by making fairly accurate selections and then to 'clear the board' one by one with quite erroneous choices.

One further point of interest concerns the children's approach to the questions of 'positive/negative identification'. It seemed clear from the children's puzzled faces (and from some of the children's comments to their parents afterwards) that the questions we were asking were considered rather 'weird' or perhaps that they were not considered 'real questions' at all. This seems quite an understandable reaction when one looks at some of the questions e.g. 'Are spoons people?', 'Are teddy-bears food?'. About half the children who were unable to answer these questions on the first occasion simply answered 'Yes' to each question. Another common response, especially with the younger children, was to repeat the question as a question or to turn it into a statement e.g. 'Eggs are people'. There is some indication also that the 'penny sometimes drops' suddenly and the children seem to see immediately that they are, in fact, being asked to answer a question which, though unusual, is also meaningful. Thus several children (14 per cent of the reliability sample) scored completely or almost completely correct records on the second testing session after having simply answered with a blanket 'Yes' on the first occasion.

It was also illuminating to watch the way in which some of the children were clearly striving to find the underlying basis for the generic concepts and vacillating in the certainty with which they approached the various tests of concept acquisition (cf. Piaget's discussion in Tanner and Inhelder, 1956, of 'equilibration' in the development of concrete operations). For example, one boy aged four years ten months making his selection of pictures of 'food' came to the 'cup' and asked himself aloud: 'This is food isn't it? We drink from it don't we?'. He then proceeded to include it as 'food'. At the same time he rejected the 'cup' as 'food' on the 'positive/negative identification test'.

In relation to the concept definition test, not surprisingly few children were able to offer clearly articulated definitions of the concepts under consideration. One might hazard a guess that 'people' would be a difficult concept for many adults to define! Seven children gave an adequate definition of 'food' mentioning that 'we eat it', but three of these failed on both parts of the test in relation to 'food'! Only one child gave anything approaching an adequate definition of 'people': 'what talk', though two mentioned 'walk' and one 'watching tele'! The majority of children simply said they didn't know or they repeated their choices by pointing to the pictures or picking them out again. Most of the definitions were offered by children over the age of four years, though two children below four, one aged three years two months and the other three years three months, offered correct definitions of 'food'. Level of verbal ability, as adjudged by the PPVT did not seem too relevant to ability to offer definitions. Thus the standardized scores of those children offering definitions for 'food' were: 140; 93; 108; 97; 82; 127; 106.

Reliability of HAPCAT items

As a rough guide to the reliability of the HAPCAT items, the test/re-test records of 17 children were examined in relation to (1) the identification of the pictures, (2) the choice of pictures in relation to the generic concepts 'food' and 'people', (3) the positive/negative identification question ('Are eggs food?') and (4) definitions.

As far as (1) is concerned, 59 per cent of the children gave the same response to the pictures on each occasion and only one child's record showed more than two changes of response from test to re-test. The majority of changes were from errors on the first test to correct responses on the re-test and many of the errors involved initial errors of recognition of the *drawings* (e.g. 'rocket' for 'crayon'). The tendency for children to identify the pictures correctly on the re-test is partly related, of course, to the fact that their errors were corrected on the first occasion i.e. they were told what the picture was meant to depict.

The pattern of responses was very similar for test (2) concerning generic concepts. About half the responses remained unchanged from the test to re-test and only three children made more than three changes (all in relation to 'people'). In general these results suggest that the children's responses remained fairly stable over a short period of time (four weeks) in relation to tests involving the identification of specific concepts as depicted in a series of pictures and in their selection of items representative of certain generic concepts. The data thus lend support to the view that these two HAPCAT items probably offer fairly reliable testing instruments.

Test (3)—the 'positive/negative identification' test showed rather less stability although even here in about two-thirds of the cases the records were the same or involved only one or two changes of response from test to re-test. In half of the cases involving changes of more than two responses the children had changed their responses from 'Yes' to every question on the first occasion to correct responses on the second occasion. The children appeared on the second testing occasion to have grasped that the questions they were being asked ('Are eggs people?', etc.), though very odd-sounding questions, were nevertheless questions. Once having accepted that the questions were deserving of answers, as it were, then they often proceeded to give correct answers.

As far as the final question (4) is concerned (definitions) again the children showed a rather less consistent response pattern from one test to the next than obtained for tests (1) and (2). The most usual change of response pattern was from a 'Don't know' to the indication of the particular pictures which the child understood to represent 'food' or 'people' either by pointing or verbalizing or both.

Teaching the generic concepts

The data were obtained from only five paired cases in which the matching process was only achieved at a very crude level and must, therefore, be treated with great circumspection. Nevertheless it was felt worthwhile to offer the results for consideration. The children's records were fairly evenly balanced, two of the 'taught' children showing an improvement over the records of the paired 'un-taught' children; two pairs of children in which the reverse situation applied and one pair in which the results showed fairly equivalent standards of performance on the two testing sessions. However, when the records of each individual child's performance were considered in detail the 'taught' children appeared to have done rather better. Thus the most dramatic improvements seen in the records related to 'taught' children'. One child of three years one month on first testing, for example, simply picked up each of the pictures from the card when asked to show the pictures of 'food' and 'people'. When re-tested he selected the 'food' pictures entirely correctly and all the 'people' with the addition of a teddy-bear. None of the 'un-taught' children showed improvements of this magnitude.

Piaget's test questions

The responses to the Piagetian questions are summarized in Table 3. The main observation

Table 3: The proportion of children answering the Piagetian questions correctly (a) in the present research, and (b) in the research of Piaget and Inhelder (1964)

	(a) *Present sample* (N=29 ages 2–4 yrs)				(b) *Piaget's sample* (N=32 for Qs. [2], [4], [5], [7] ages 6–9 yrs.; N=20 for Qs. [3] and [6] ages 5–6 yrs.)
	On first testing		*After demonstration*		
Q	*No. answering correctly*	%	*No. answering correctly*	%	% *answering correctly*
1	29	100	—	—	not asked
2	28	97	29	100	47
3	26	90	28	97	71
4	29	100	—	—	21
5	20	69	24	83	81
6	13	45	25	86	71
7	19	66	24	83	33

which stands out from the table is that there is little correspondence between the proportion of children answering the questions accurately in Piaget's samples (ages 5-9) and the children in the present sample (ages two to four). This lack of correspondence will be discussed later. There was, similarly, very little apparent correspondence between the difficulty levels of the different questions. Again we will return to this point.

Over 80 per cent of the children tested in the present research answered the questions correctly either the first time or after a 'concrete' demonstration of the problem using the picture cards in front of them. The question which posed most problems for the older children in Piaget's research (*viz.* 'Are all the tulips [or roses] some of the flowers?') formulated in the present research as 'Are all the ladies some of the people?', Question (4) caused *no* problems at all to the children of two to four years of age in our sample. They all answered correctly 'Yes'. On the other hand, whereas 81 per cent of Piaget's sample answered correctly the logically identical question (5) only 69 of our sample did so. The question the children found most difficult to answer in our research was question (6): 'If you take away all the people will there be any ladies left?'. After a concrete demonstration of the problem using the pictures, however, 86 per cent of the children also answered this correctly.

The individual records of the children also revealed some interesting data relevant to question (7): 'Are there more ladies or more people?'. We have already argued that this seems rather like a trick question. It seems as if the child is led to suppose by the phrasing of the question that the questioner has created an artificial distinction between 'ladies' and 'people' for the purpose of the test. In answering, the child appears to accept the artificial dichotomy suggested to him by the authority figure of the tester. For example, one boy aged three years nine months having answered this question correctly by saying 'more people', then went on to give a spontaneous explanation of his reasoning. 'There are three ladies and three people'—here appearing to follow the lead given by the questioner. Then, as if dismissing the 'trick dichotomy' and adopting his own 'alternative' interpretation, he carried on: 'six people if you count them altogether . . . if they're altogether there are more people'.

Seven out of the 29 children (or 24 per cent) answered *all* the questions correctly and a further seven children (i.e. a total of nearly half the sample altogether) having failed on *one* question—usually question (6)—answered all correctly after one demonstration. Four of the 29 children got all the Piagetian questions right but failed to complete the HAPCAT tests at all accurately. For example, one girl, aged three years ten months, scored 100 per cent on the questions but was unable to pick out the pictures of people from amongst the distractors. She chose the people together with 'dog', 'cat', 'bird' and 'teddy-bear', making similar errors in the 'positive/negative identification' test. Another illustration is of a boy aged four years ten months who selected the items correctly as representing 'people' but made errors in his 'positive/negative identification', including 'bird' and 'dog' as people. Some children also showed records in the opposite direction i.e. they scored accurately on the HAPCAT items but did not complete all the Piagetian questions correctly. Five children, in fact, gave accurate selections of both 'food' and 'people' and were also able to answer the follow-up identification questions accurately but made one error on the Piagetian questions (Question [6] in four out of five of the cases). At the same time, with one exception, all these children were able to give a correct answer to the question after using the picture cards to work out the problem. Four children (i.e. 14 per cent of the sample of 29 given the Piagetian questions) gained full scores on all the HAPCAT items and also answered all the Piagetian questions correctly. All these children had PPVT scores in the 'above average' range (the lowest standardized score as 123). Their ages were three years two months, four years six months, four years seven months and four years seven months.

The children's responses to the Piagetian questions (unlike those to the HAPCAT items) appeared comparatively unrelated to chronological age, PPVT score and socioeconomic background. Comparing the performance of children on the Piagetian tests (scores of five plus *versus* four or less) with chronological age ('four' *versus* 'under four') the chi-squared value was 0·02; the equivalent chi-squared result for comparison of the Piagetian tests with the PPVT scores was 0·02. Similarly, socioeconomic background appeared unrelated to performance on the Piagetian tests when tested by a 3×2 chi-squared analysis (socioeconomic groups I, II and III, as opposed to scores of 5+/4 or less on the Piagetian questions).

Discussion

Although the research sample was not very large (N=56) it was nevertheless reasonably representative of the characteristics normally found in much larger samples of children in relation to certain important variables (e.g. the socioeconomic backgrounds of the children and their verbal ability scores) though not in the proportion of the sexes (boys outnumbering girls in the present sample). It seems justifiable to postulate, therefore, that the results may well relate more generally to pre-school children in urban environments in England, though further corroborative work would be required to substantiate this claim.

As far as the first question is concerned ('To what extent had the children acquired certain specific [i.e. non-generic] concepts?') the evidence suggests that these children had formed many specific concepts as adjudged by their ability to identify the pictures of certain objects from amongst a number of similar object concepts (e.g. 'dog' from 'cat', 'cup' from 'spoon') when presented in an entirely novel setting. Thus they appeared to have abstracted the common features of certain objects (e.g. 'cups') and have demonstrated an ability to generalize their awareness of the concept to new situations in which it was necessary to recognize a novel presentation of the object and to distinguish one concept ('cup') from another similar concept ('spoon').

Similarly, with reference to our second question ('To what extent were these pre-school children capable of showing an understanding of generic or class concepts?') the evidence suggests that many children in this age group have formed some generic concepts and there is an indication that there may be something to be said for further experiments on the *teaching* of generic concepts at the pre-school level. Judged by very strict standards (correct answers to all the HAPCAT items and to all the Piagetian questions involving 'class inclusion') four out of 29 children (i.e. 14 per cent of the sub-sample of children given the Piagetian questions in addition to the HAPCAT items) showed what appeared to be an 'unshakeable' grasp of class inclusion. When one applies different test standards to the performance of the children, their grasp of class inclusion appears either more or less certain. For example, if one takes Piaget's test questions alone, 24 per cent of the sub-sample (seven out of 29 children) could be said to have attained an understanding of class inclusion, and this percentage rises to 80 per cent if one allows the scores of children who gained completely accurate answers after being allowed to work out the problem using the card pictures. On the other hand, if one takes only performance on the HAPCAT items as the criterion for the establishment of class inclusion, one finds rather different proportions.

The two HAPCAT items mainly relevant to the generic concept involved (a) the selection of pictures representative of the concept and (b) the positive/negative identification of the concept ('Are spoons food?' etc.). Between a quarter and a third of the pre-school children under consideration had attained an understanding of the generic concept if one takes as the criterion completely accurate performance on both these tests. On the other hand, if one takes accurate performance of the first test (a) only then the percentages rise to about half the sample of 56 children (59 per cent choosing the pictures of 'people' correctly and 43 per cent the pictures of 'food').

In general, therefore, the present results suggest that quite a substantial proportion of children in the age group two to four years (i.e. the age group normally taken to be in the pre-conceptual stage of development) are, in fact, quite capable of forming not only specific concepts but also, as Welch (1940) maintained, hierarchical classification involving class inclusion. The number of generic concepts which find a place in the vocabulary of the pre-school child depends partly, of course, upon the language used by adults when talking to children, a point very clearly argued by Brown (1958). Thus 'cutlery' is probably a generic term which finds little usage whereas 'clothes' is frequently used in conversation between adults and children. Similarly, the extent to which a child is able to define the distinguishing attributes of any particular concept will vary with general experience, and intellectual maturity, just as adults constantly refine their own understanding of various concepts in the light of new experiences and increasing knowledge.

It seems clear from the results, however, that many pre-school children are able to understand the essential characteristics of a number of generic concepts. If these results are replicated in other studies they would suggest that some modification is required in the standard 'Piagetian texts' which claim that children between two to four years cannot form generic concepts.

The key point of dissension between these results and those of Piaget and Inhelder (1964) does not refer to the Piagetian thesis concerning the relationship between operational thinking and hierarchical classification (though it does cast doubt upon the suggested age level at which hierarchical classification can be achieved). The dissension rather relates to the *criterion tests* which are accepted as adequate for the demonstration of a child's acquisition of hierarchical classification involving class inclusion. Thus Piaget claims that his questions provide a sensitive and reliable test for the acquisition of class inclusion but on the evidence of the present research findings we would have to doubt this assertion.

The Piagetian questions, in fact, showed very little discriminatory power between the results of children of different age and ability levels. Unlike the HAPCAT items the Piagetian questions appeared unrelated to ability and also to chronological age at this level (two to four years), though the results of Piaget and Inhelder (1964) do show some age/concept development relationship with a sample of children between five and ten years. Similarly, the test question which Piaget's subjects found most difficult was answered correctly by each of the 29 children to whom we presented the item. This was the question 'Are all the ladies some of the people?' Further, whereas, in Piaget's sample, a *higher* proportion of children were able to answer the question containing a logically identical element (in our research 'Are all the people ladies or are some of the people ladies?') the present sample found this question *more* difficult. It seems probable that these discrepancies probably relate to the age ranges of the children involved. Thus the type of test questions used by Piaget to demonstrate the attainment of an understanding of generic concepts may, in fact, be quite inappropriate for use with children below the age of four years. For example, four of the children who answered all the Piagetian questions correctly were unable to answer the HAPCAT items on generic concepts accurately. Thus it may be that one is simply demonstrating that children at this age are not yet capable of appreciating the unusual linguistic structure of the question rather than that they have not attained an understanding of generic concepts. The 100 per cent 'Yes' to the question 'Are all the ladies some of the people?' in our sample of two- to four-year-old children may simply represent an 'uncomprehending affirmative' in the same way that children who did not understand the odd format of the HAPCAT 'positive/negative identification' questioning ('Are eggs people?' etc.) simply answered with a blanket 'Yes'.

It is possible that the Piagetian tests may show something about the level of linguistic comprehension when used with pre-school children but they do not appear to be useful guides to an understanding of hierarchical classification at this age. With older children they may perhaps help to demonstrate the degree of certainty with which the child has attained some generic concepts but it is also seemingly possible that they could be rather more closely related to the child's ability to answer questions concerning class inclusion which are phrased in the manner of verbal reasoning 'teasers'.

The HAPCAT items, on the other hand, do seem to provide a fairly meaningful measure of a child's understanding of generic concepts. The items appear, on this initial testing, to offer a reasonable degree of reliability (an aspect which would be interesting to check in relation to the Piagetian questions, though we did not do so on this occasion). They also showed a good discriminatory power in relation to the age and ability levels of the children, the higher the age and ability level the higher the scores on the items ($P<0{\cdot}01$ in relation to verbal ability and $P<0{\cdot}001$ in relation to age). There seems to be a substantial case, therefore, for arguing that children are able to respond appropriately to test items involving an understanding of generic concepts provided that the items are presented in a format which is meaningful to them. The HAPCAT items appear to satisfy such criteria.

References

BEARD, R. M. (1969). *An Outline of Piaget's Developmental Psychology*. London: Routledge & Kegan Paul.

BLANK, M. (1968). 'Experimental approaches to concept development in young children', in LUNZER, E. A. and MORRIS, J. F. (Eds.) *Development in Human Learning*. London: Staples.

BROWN, R. W. (1958). 'How shall a thing be called?' *Psychol. Rev.*, **65**, 1, 14–21.

BROWN, R. W. (1967). 'Language and categories', in BRUNER, J. S., GOODNOW, J. J. and AUSTIN, G. A. *A Study of Thinking*. New York: J. Wiley & Sons.

DUNN, L. M. (1965). *Expanded Manual for the Peabody Picture Vocabulary Test*, American Guidance Services Inc.

MOSS, J. W. and EDWARDS, P. (1960). 'The PPVT with English children', *Brit. J. Educ. Psychol.*, **30**, 82.

PIAGET, J. (1962). 'The stages of the intellectual development of the child', in WASON, P. C. and JOHNSON-LAIRD, P. N. (Eds.) (1968). *Thinking and Reasoning*. Harmondsworth: Penguin.

PIAGET, J. and INHELDER, B. (1964). *The Early Growth of Logic in the Child: Classification and Seriation*. London: Routledge & Kegan Paul.

RUSSELL, D. H. (1956). *Children's Thinking*. London: Ginn.

STONES, E. (1966). *An Introduction to Educational Psychology*. London: Methuen.

TANNER, J. M. and INHELDER, B. (Eds.) (1970). *Discussions on Child Development, Vol. 1*. London: Tavistock Publications Ltd.

WELCH, L. (1940). 'A preliminary investigation of some aspects of the hierarchical development of concepts', *J. Gen. Psychol.*, **22**, 359–78.

British Journal of Educational Psychology

Prof. N. J. ENTWISTLE, Department of Educational Research, University of Lancaster

Vol. 45 Part 2 June 1975

KALLOS, D. and LUNDGREN, U. P. Educational Psychology: Its Scope and Limits.

BOYDELL, D. Pupil Behaviour in Junior Classrooms.

TROWN, E. A. and LEITH, G. O. M. Decision Rules for Teaching Strategies in Primary Schools: Personality-Treatment Interactions.

ROBINSON, W. P. Boredom at School.

GAYNOR, J. and MILLHAM, J. Individual Differences and Course Performance.

STOTT, D. H. and ALBIN, J. B. Confirmation of a General Factor of Effectiveness-Motivation by Individual Tests.

BROADBENT, D. E. Cognitive Psychology and Education.

PEEL, E. A. Predilection for Generalising and Abstracting.

KELLY, M. and PHILP, H. Language and Cognitive Development: Vernacular Test Instructions in Relation to Task Behaviour among Highlands Children of Papua, New Guinea.

VALENTINE, E. R. Performance on two Reasoning Tasks in Relation to Intelligence, Divergence and Interference Proneness.

CHOPPIN, B. Guessing the Answer on Objective Tests.

MUICHERJEE, B. N. The Factorial Structure of Wechsler's Pre-School and Primary Scale of Intelligence at Successive Age Levels.

RESEARCH NOTES:

BANRETI-FUCHS, K. M. Attitudinal, Situational and Mental Health Correlates of Academic Achievement at the Undergraduate University Level.

ROWELL, J. A. and RENNER, V. J. Personality, Mode of Assessment and Student Achievement.

BOOK REVIEWS.

PUBLICATIONS RECEIVED.

Scottish Academic Press Limited, 25 Perth Street, Edinburgh EH3 5DW, Scotland, UK

Primary French in the Balance*

This article reviews the evidence presented in the final report of the NFER French Project (1964-74) published last year under the same title.

Clare Burstall, Deputy Director, NFER

Summary

During 1963, arrangements were made by the then Ministry of Education for a national experiment in foreign-language teaching to be carried out in selected primary schools in England and Wales. The main purpose of the experiment, which came to be known as the Pilot Scheme for the teaching of French in primary schools, was to discover whether it would be both feasible and educationally desirable to extend the teaching of a foreign language to pupils who represented a wider range of age and ability than those to whom foreign languages had traditionally been taught. Under the Pilot Scheme, French was to be introduced into the primary school curriculum on an experimental basis from September 1964 onwards. The choice of French as the language to be taught was virtually inevitable, since it would have been impossible to provide an adequate teaching force for the implementation of the experiment if any language other than French had been chosen. In most of the schools taking part in the Pilot Scheme, French was to be taught throughout the primary stage of the experiment by class teachers who had received special in-service training, rather than by specialist teachers of French. Arrangements were made to provide continuity of teaching at the secondary stage, so that all the pupils taking part in the experiment would be able to continue learning French without interruption for at least five years.

Once the experiment had been set up, it was agreed that its effects should be evaluated over a period of years by the NFER. In the event, the NFER evaluation spanned the period 1964–1974, taking the form of a longitudinal study of three age-groups or 'cohorts' of pupils attending the schools taking part in the experiment. The sole criterion on which pupils were chosen for inclusion in one of the experimental cohorts was their date of birth. In the first instance, French was to be taught to all eight-year-old pupils in the selected primary schools from September 1964 onwards; thereafter, the teaching of French was to be extended to a further year-group each autumn, until it involved all pupils in the 8-11 age-range. Thus, the first cohort to come under study was composed of all pupils in the large primary schools taking part in the experiment who fell within the age-range 8·0–8·11 on 1st September 1964 and all pupils in the small primary schools who fell within the age-range 8·0–9·11 on that date: this provided a sample of approximately 5.700 pupils. (A wider age-range was sampled in the small primary schools, to avoid the creation of unworkably small groups.) The second cohort was composed of all pupils in the large primary schools in the sample who fell within the age-range 8·0–8·11 on 1st September 1965: this provided a sample of approximately 5,300 pupils. Pupils in the small primary schools were not represented in the second cohort, since most children of an appropriate age had already been included in the French classes set up for the first cohort.

Originally, the NFER evaluation was to have been based entirely on a longitudinal study of the pupils forming the first two experimental cohorts. It was hoped that the results of this study would provide sufficient information to enable valid conclusions to be drawn regarding the feasibility and advisability of teaching French at the primary level. As the experiment progressed

---

*Offprints are available from the NFER Publishing Company, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS, as follows: £1·50 per 10; £2·75 per 20; and £6·50 per 50 copies, postage included. Payment *must* accompany order.

however, it became clear that the pioneer status of the first cohort had entailed an atypical introduction to French. During the first year of the Pilot Scheme, for instance, difficult staffing problems were encountered which had not always been foreseen: in some primary schools, French teachers were absent without replacement for a whole term in order to attend intensive language courses in France; in others, no trained staff were available to teach French during the first term of the experiment, with the result that the first cohort pupils in these schools started to learn French one term later than the others in their age-group. The first year of the experiment could therefore be regarded with some justification as an essentially exploratory period, calling into question the validity of using the results obtained from the study of the first cohort as a basis for future comparison. In view of these circumstances, it was considered advisable to extend the evaluation to a third cohort of pupils: those who would begin their study of French in September 1968. The third and final cohort to come under study was thus composed of all pupils in the large primary schools still taking part in the experiment who fell within the age-range 8·0–8·11 on 1st September 1968 and all pupils in the small primary schools who fell within the age-range 8·0–9·11 on the same date: this provided a sample of approximately 6,000 pupils. Inclusion in the experimental sample was again determined solely by the age of the pupil. This meant that the sample was drawn from all the socioeconomic strata normally represented within the national educational system and, in consequence, was characterized by a wide range of ability.

The time-span of the evaluation did not allow all the pupils taking part in the experiment to be studied for an equal period of time. The pupils in the first and third cohorts were under study for a total of five years: three years in the primary school and two years in the secondary school. The pupils in the second cohort were under study for a longer period: three years in the primary school and five years in the secondary school. During the ten-year period of the evaluation, the main aims of the study were: (i) to investigate the long-term development of pupils' attitudes towards foreign-language learning; (ii) to discover whether pupils' levels of achievement in French were related to their attitudes towards foreign-language learning; (iii) to examine the effect of certain pupil variables (such as age, sex, socioeconomic status, perception of parental encouragement, employment expectations, contact with France, etc.) on level of achievement in French and attitude towards foreign language learning; (iv) to investigate whether teachers' attitudes and expectations significantly affected the attitudes and achievement of their pupils; (v) to investigate whether the early introduction of French had a significant effect on achievement in other areas of the primary school curriculum.

The main findings to emerge during the earlier years of the evaluation were published in two interim reports (Burstall, 1968, 1970); the recent publication of the final report (Burstall *et al.*, 1974) brought together both the earlier and the later findings and provided an overall view of the effects of the experiment during both its primary and secondary stages. What follows is an attempt to review briefly the research evidence presented in the final report, but it must be borne in mind that limitations of space will inevitably impose a certain selectivity on this review.

Findings

1. *An optimum age for foreign language learning?*

The belief that young children are better-equipped than older children or adults to learn foreign languages with speed and efficiency underlies the recent expansion of foreign language teaching at kindergarten and elementary school level. Much of the impetus for the early introduction of foreign language study can be attributed to the influence of Wilder Penfield's work. Penfield has frequently reiterated the view that the young child's brain is uniquely well-adapted for language learning and that there is an 'optimum age' during which 'multiple languages may be learned perfectly, with little effort and without physiological confusion' (Penfield, 1953). Penfield situates this 'optimum age' for language learning within the first decade of life, after which period a 'built-in biological clock' inexorably records the lost educational opportunity (Penfield, 1964). Further attempts to acquire foreign languages will be crowned with only modest success. The age of the learner is thus the most critical factor in the language learning process: if it is to be successful, foreign language learning must take place before the age of ten, 'in accordance with the demands of brain physiology' (Penfield and Roberts, 1959).

The findings of the evaluation are not in harmony with this view. Other things being equal, the older children tended to learn French more efficiently than the younger ones did. Pupils taught French from the age of eight did not show any substantial gains in achievement, compared with those who had been taught French from the age of eleven. By the age of 16, the only area in which the pupils taught French from the age of eight consistently showed any superiority was that of listening comprehension. Even there, the differences between the various groups of pupils, although statistically significant, were hardly of a substantial nature, being of the order of two to four points difference on a 28-item test—a fairly minimal return for the extra years spent learning French in the primary school.

Where the pupils taught French in the primary school did appear to gain was in attitude rather than in achievement. When they were successful in their efforts to learn French, they maintained a more favourable attitude towards speaking the language than did those who were not introduced to French until the age of eleven. However, this more positive attitude was not reflected in a correspondingly higher level of achievement. The most conservative interpretation which the available evidence would appear to permit is that the achievement of skill in a foreign language is primarily a function of the amount of time spent studying that language, but is also affected by the age of the learner, older learners tending to be more efficient than younger ones. Penfield's contention that the first ten years of life constitute a 'critical period' for foreign language learning remains unsupported by direct experimental evidence.

2. *Sex differences*

Throughout the period of the experiment, the girls in the experimental sample scored significantly higher than the boys on all tests measuring achievement in French. From the age of 13 onwards, the low-achieving boys in the sample tended to 'drop' French to a significantly greater extent than the low-achieving girls did, with the result that the sample of girls still learning French at the age of 16 represented a considerably wider range of ability than the corresponding sample of boys did. However, in spite of this disparity, the girls still continued to score significantly higher than the boys on each of the French tests. If the composition of the original experimental sample had been maintained until the end of the secondary stage of the experiment, there can be little doubt that sex differences in achievement in French favouring the girls would have been even more prominent.

The attitudes of the girls towards learning French were also consistently more favourable than those of the boys. This finding is similar to that reported by Nisbet and Welsh (1972) in their evaluation of the teaching of French in schools in Aberdeen. Differences in attitude undoubtedly stem partly from the different employment expectations of boys and girls and the extent to which foreign language learning can be perceived as relevant to their occupational requirements. Other studies have reported that adolescent pupils of both sexes tend to view the enhancement of vocational success as the primary function of education and, in consequence, place a high value on school subjects, such as English and Mathematics, which have an obvious relevance to their future employment prospects. Girls and their parents are also reported to accept the vocational value of foreign language learning, whereas boys and their parents are reported not to do so (Schools Council, 1968; Sumner and Warburton, 1972). Where achievement in a particular area of the curriculum can be shown to vary systematically with the sex of the pupil, it seems highly probable that parents are transmitting to their children the accepted values of the wider society. The findings of both the primary and the secondary stages of the evaluation would support this view.

3. *Socioeconomic factors*

During the early stages of the evaluation, it rapidly became apparent that the pattern of results emerging from the inquiry pointed to a linear relation between the pupil's socioeconomic status and his level of achievement in French. On each occasion of testing, high mean scores tended to coincide with high-status parental occupation and low mean scores with low-status parental occupation: this was equally true for both boys and girls and for the pupils in each of the experimental cohorts. This association between socioeconomic status and level of achievement in French became even more prominent during the early years of the secondary stage of the experiment, when the pupils in the experimental sample were distributed unevenly among the different types of secondary school. Disparities in the socioeconomic composition of the different types of school were closely paralleled by significant differences in the pupils' level of achievement in French: the highest level of achievement was reached by the grammar school pupils, the lowest by the secondary modern and bilateral school pupils, while the comprehensive school pupils occupied an intermediate position.

Similarly, during the early years of the secondary stage of the evaluation, there was a positive tendency for the percentage of pupils with favourable attitudes towards learning French to increase with social status. This tendency was clearly illustrated by significant differences in attitude between pupils attending the different types of secondary school: whatever the specific point at issue, favourable attitudes were most characteristic of the grammar school pupils in the experimental sample and unfavourable attitudes of the secondary modern school pupils, particularly of the boys. The secondary modern school boys were characterized by the growing conviction that they were unlikely ever to speak French in after-school life, linked with the diminishing belief that their parents were in favour of their continued study of French.

After their second year in secondary school, the secondary modern school boys 'voted with their feet', 'dropping' French in such large numbers that, by the end of the secondary stage of the experiment, the tendency for the percentage of pupils with favourable attitudes towards learning French to increase with social status was barely discernible. By that stage, the attitudes of the pupils attending the different types of secondary school were remarkably uniform, reflecting the restricted socioeconomic range represented in the attenuated experimental sample.

4. *Achievement in the small schools*

The primary schools involved in the experiment had a wide geographical distribution and varied greatly in size. The experimental sample included a number of small rural schools, the smallest of which had 16 pupils on roll, the largest 160 pupils. Throughout the primary stage of the experiment, the pupils in these small rural schools consistently maintained a higher level of achievement in French than did those in the larger urban schools. This finding was an unexpected one. The classes in the small schools usually contained pupils differing widely in age and ability, and classroom conditions were often inimical to the teaching of French by audio-visual means. In spite of these apparently adverse circumstances, however, the test performance of the pupils in the small schools was consistently superior to that of the pupils in the large schools. At first it was thought possible that this superiority might be partly attributable to the small size of the classes in which the pupils attending the small schools were taught French, but this did not prove to be the case: further analysis of the data revealed no association between size of class and level of achievement in French. The possibility remains that the high level of achievement in the small schools stems partly from the heterogeneous nature of their classes. If a given class contains pupils who vary greatly in age and ability, the individual pupil is not in direct competition with others of his own age-group: the concept of a 'standard' of achievement, which a pupil of a given age 'ought' to be able to reach, is difficult for either teacher or pupil to acquire. The classroom situation in the small school tends to encourage co-operative behaviour and to lack the negative motivational characteristics of the competitive classroom in which success for a few can only be achieved at the expense of failure for many.

The differences in achievement in French between pupils in small primary schools and those in large primary schools proved to be of an extremely persistent nature. A follow-up study of the pupils who had formerly attended small primary schools showed that, even after two years in the secondary school, these pupils continued to achieve significantly higher scores on the French tests than did their classmates who had formerly attended large primary schools. There is some evidence (Dale, 1969, 1971; Sumner and Warburton, 1972) that pupils in small schools tend to form closer relationships with their teachers than pupils in large schools do and, in addition, are more responsive to evidence of 'teaching effort'. The higher level of achievement in French of the pupils in the small schools may thus be at least partly attributable to the early establishment of good teacher-pupil relationships and the subsequent development of positive attitudes towards further learning.

5. *Achievement in other areas of the primary school curriculum*

Since the introduction of French into the primary school curriculum inevitably reduced the time available for other school activities, the possibility arose that this reduction in time might be sufficient to exert a detrimental effect on the experimental pupils' acquisition of the basic skills of literacy and numeracy. It was equally possible that the stimulus of learning a foreign language might serve to increase the pupils' understanding of their own language and so exert a generally beneficial influence on the development of their skills. In order to investigate these possibilities, it was necessary to assess the experimental pupils' level of attainment in the 'basic' areas of the curriculum before they began to learn French and to repeat this procedure three years later, at the end of the primary stage of the experiment, comparing their test performance with that of control pupils who were not exposed to any foreign language teaching during their period in primary school. This assessment was carried out by means of a battery of general attainment tests, designed to measure both verbal and mathematical skills.

The results of the general attainment survey revealed that the introduction of French had not exerted any significant influence on achievement in other areas of the primary school curriculum. Variations in test performance were always accompanied by corresponding variations in the social composition of the groups concerned: there were no indications of a major trend in either a positive or a negative direction. It must be emphasized, of course, that these findings are specific to the tests used in the evaluation and cannot be generalized to areas of the curriculum which fall outside their range, but they are, nevertheless, in harmony with findings reported over the years by a number of other investigators.

6. *Attitudes and achievement in co-educational and single-sex schools*

The findings of the secondary stage of the evaluation largely support those reported by Dale (1974) in his study of attitudes and achievement in co-educational and single-sex schools. Dale found that boys in single-sex schools tended to have a more favourable attitude towards learning French and to reach a higher level of achievement in French than did boys in co-educational schools. He also found some evidence of a reverse trend for girls, but differences in attitude and achievement were less marked for girls than they were for boys. However, Dale did note a tendency for girls in co-educational schools to show a greater reluctance to 'speak out in class'.

With regard to achievement in French, the findings of the secondary stage of the evaluation are clear-cut: both boys and girls in single-sex schools reached a higher level of achievement in French than did pupils of either sex in co-educational schools. With regard to pupils' attitudes towards learning French, the findings indicate greater differences for boys than for girls: boys in single-sex schools had markedly more favourable attitudes towards learning French than had boys in co-educational schools. It is interesting to note, however, that a significantly higher percentage of boys in co-educational than in single-sex schools enjoyed being at school. Dale had also found that boys in co-educational schools expressed greater 'happiness' in school than did those in single-sex schools.

Girls in single-sex schools showed a slight tendency to adopt a less favourable attitude towards learning French than did girls in co-educational schools, but girls in co-educational schools were significantly more fearful about speaking French in class. This latter finding is similar to Dale's evidence regarding the greater reluctance of girls in co-educational schools to 'speak out in class' and is no doubt attributable to the stresses of adolescent sexuality. There were no sex differences in pupils' attitudes towards speaking French in class during the primary stage of the experiment, even though the majority of the pupils were in co-educational schools. It was only during the secondary stage of the experiment that the girls revealed an ever-growing anxiety about speaking French in front of other members of the class.

Limitations of space prevent a fuller review of the evidence, but the interested reader may care to note that the final report also contains a wealth of teachers' and pupils' comments on various aspects of the experiment, some evidence regarding regional differences in attitude, the effect of teachers' contact with France, the influence of teachers' attitudes on pupils' achievement in French, and the impact of early success or failure on later achievement in French and attitudes towards learning the language.

References

BURSTALL, C. (1968). *French From Eight: A National Experiment*. Slough: NFER.

BURSTALL, C. (1970). *French in the Primary School: Attitudes and Achievement*. Slough: NFER.

BURSTALL, C., JAMIESON, M., COHEN, S. and HARGREAVES, M. (1974). *Primary French in the Balance*. Slough: NFER.

DALE, R. R. (1969). *Mixed or Single-Sex School? Vol. 1: A research study in pupil–teacher relationships*. London: Routledge and Kegan Paul.

DALE, R. R. (1971). *Mixed or Single-Sex School? Vol. 2: Some social aspects*. London: Routledge and Kegan Paul.

DALE, R. R. (1974). *Mixed or Single-Sex School? Vol. 3: Attainment, attitudes and overview*. London: Routledge and Kegan Paul.

NISBET, J. D. and WELSH, J. (1972). 'A local evaluation of primary school French', *Journal of Curriculum Studies*, **4**, 2, 169–175.

PENFIELD, W. (1953). 'A consideration of the neurophysiological mechanisms of speech and some educational consequences', *Proceedings of the American Academy of Arts and Sciences*, **82**, 201–214.

PENFIELD, W. (1964). 'The uncommitted cortex: the child's changing brain', *Atlantic Monthly*, **214**, 1, 77–81.

PENFIELD, W. and ROBERTS, L. (1959). *Speech and Brain-Mechanisms*. Princeton, New Jersey: Princeton University Press.

SCHOOLS COUNCIL (1968). *Young School Leavers*, Inquiry 1. London: HMSO.

SUMNER, R. and WARBURTON, F. W. (1972). *Achievements in Secondary School*. Slough: NFER.

Educational Research Volume 17 Number 3

Personality, Ability and Achievement in Primary School Children

Kathleen Orme, Abbeydale Middle School, Sheffield

Summary

This study investigated the relationship of personality, ability and school achievement. The findings here were that:

(1) intelligence is the major determinant of school achievement;

(2) bright children tend to come from smaller families than dull children;

(3) relatively unstable children have a better level of school achievement than stable children;

(4) extraversion–introversion had no effect on school achievement.

Introduction

In recent years, attention has been directed to the possible importance of personality variables in school achievement (e.g. Eysenck and Cookson, 1969a and b; Elliott, 1972; Entwistle, 1972). The findings are complex but suggest that in the early school years stable extraverts are the most successful. In later school life and at college and university unstable introverts appear to be the most successful. As these links between personality and attainment derive from studies of very large numbers or of specially selected samples, they may be only of theoretical interest. This study is of a relatively small but unselected group of children and investigates the practical value of intelligence and personality tests in predicting attainment.

Method and Results

The study was of an entire year (M3) of a Sheffield primary school whose children mostly came from a large, adjoining pre-war council estate. The year was split into four classes, graded according to school progress. The psychological tests used were the Junior Eysenck Personality Inventory (JEPI), which measures neuroticism and extraversion-introversion (Eysenck, 1965), and the Coloured Progressive Matrices (CPM), a non-verbal intelligence test (Raven, 1965). Grades (A to E) of school performance were available in mathematics, English, environmental studies and science, together with a Reading Quotient.

The 112 children who completed the JEPI had a mean neuroticism score of 16·74 (SD 3·67) and a mean extraversion score of 13·83 (SD 4·56). They are, therefore, rather more neurotic and introverted than the test's norms for this age. This is of some interest in that the normative sample were mostly from Rotherham, near to the present children's neighbourhood. There were no sex or class differences in personality test scores.

Of these 112 children, 109 completed the CPM on a subsequent day, obtaining a mean

Table 1: Score characteristics of extreme groups

Neuroticism	Stable group (S) N = 25 Score < 15	Unstable group (U) N = 24 Score > 19
Extraversion	Introvert group (I) N = 24 Score < 11	Extravert group (E) N = 22 Score > 17
Intelligence	Dull group (D) N = 24 Score < 25	Bright group N = 20 Score > 33

score of 28·61 (SD 5·74). The test's norms give 28 as 100 IQ for this age (11 years), indicating the present sample to be a representative one for intelligence. There were no sex differences but mean scores for Classes 1, 2, 3 and 4 were 32·07, 30·55, 27·42 and 24·93 respectively. There was no relationship between JEPI and CPM scores.

For neuroticism, extraversion and intelligence, the highest and lowest scoring quarters were identified. Table 1 summarizes the cut-off test performance scores that were finally used. The numbers in the resulting groups are not equal, nor are they exactly quarters of the total sample. Whatever test score is used, the resulting groups will never be exactly similar to the contrasting group, nor exactly one-quarter of the total.

Table 2 shows the bright and dull groups are differently distributed among the four classes at a statistically significant level. Table 3 goes on to summarize the school grading data for the bright and dull groups. The various gradings have to be added together in some way to make statistical treatment feasible. It was decided that A+B V C+D+E would be the most useful as it would contrast the good achievers against the average and poor achievers. For all four attainments, the bright group does significantly better than the dull group. Ten children had a reading quotient of less than 90. Nine of these belonged to the dull group, again a statistically significant difference (chi^2=5·87, df^1, Yates' correction, P<0·02).

Table 2: Bright and dull groups and class membership

	Class 1	2	3	4
Bright	12	6	1	1
Dull	0	3	8	13

chi-square, bright versus dull, Classes 1+2 v. 3+4 =23·22 (P<0·0005, df^1, Yates' correction)

There were no significant relationships between any of the school achievement data and extraversion. There were significant relationships between gradings in mathematics, English and science and neuroticism (see Table 4). In each case, the more unstable had the better grades, the same trend (but non-significantly) occurring with regard to environmental studies and reading.

Table 3: School gradings for bright and dull groups

		A	B	C	D	E
Mathematics	Bright	7	8	4	1	0
	Dull	0	0	12	7	2
English	Bright	7	9	4	0	0
	Dull	0	0	15	3	3
Environmental Studies	Bright	5	4	11	0	0
	Dull	0	0	13	7	1
Science	Bright	3	8	9	0	0
	Dull	0	1	7	6	1

chi-square, bright versus dull, Grades A+B v. C+D+E (all df^1, Yates' correction)

Mathematics	21·69	(P<0·0005)
English	23·52	(P<0·0005)
Environmental Studies	9·62	(P<0·0005)
Science	6·86	(P<0·01)

Table 4: School gradings for the stable and unstable groups

		A	B	C	D	E
Mathematics	Stable	0	2	13	6	1
	Unstable	3	9	6	3	2
English	Stable	0	4	13	4	1
	Unstable	5	8	7	0	3
Environmental Studies	Stable	0	1	11	9	0
	Unstable	0	7	13	2	1
Science	Stable	0	3	8	8	0
	Unstable	0	10	5	3	1

chi-square, stable versus unstable, Grades A+B v. C+D+E (all df^1, Yates' correction)

Mathematics	7.69	(P<0·01)
English	5·49	(P<0·02)
Environmental Studies	3·30	(NS)
Science	4·20	(P<0·05)

Table 5: Intelligence and family size

	1, 2 or 3 children	*4 or more children*
Bright group	15	5
Dull group	6	18

chi-square = 7·38 (P<0·01, df^1, Yates' correction)

No significant relationships were found with regard to personality variables and family size or position. Table 5 indicates a significant relationship between ability and family size.

Discussion

The present results suggest that in moderately-sized samples, the relationship of extraversion to school achievement can be discounted. It is of interest that the more unstable do better than the stable. Such a finding, which tends to recur in varying forms throughout the literature, suggests moderate levels of anxiety tend to facilitate learning. As such, it may support the Yerkes-Dodson law that postulates a curvilinear relationship between anxiety and learning.

The most significant feature in the present study, however, is the importance of intellectual level in determining school performance. A child's level of ability (measured here on a test relatively independent of attainments and probably reflecting intellectual endowment) was closely related to all measures of school achievement and to particular class membership. Bright children tend to come from smaller families than dull children. This finding is again in line with the general literature reporting a negative relationship between ability and family size.

These results suggest that the major determinant of school achievement is ability. At the individual school level, extraversion is an unimportant variable, but there is some evidence that moderate amounts of instability contribute to success. It is assumed in this respect, that the unstable group here can be considered the top end of the normal range and that severely unstable children are very few in an ordinary school setting. Personality variables may be shown to be of more statistical (and theoretical) interest with large scale surveys. But in considering an individual child in a particular school, the present results suggest personality is of little predictive value in estimating a child's school achievement.

There is some evidence in the literature that up to about 12 years, extraversion is decreasingly related to educational achievement. From 12 years on, introversion is increasingly related to attainment. It has been suggested that this phenomenon reflects the varying effects of task complexity on the relationship between extraversion-introversion and attainment. It is possible to argue that the present sample, being near to 12 years in age, showed an absence of a relationship between attainment and extraversion-introversion because of these reasons. Even so, an examination of the literature shows that correlation coefficients between extraversion-introversion and attainment rarely exceed 0·20. At the most, the common variance is, therefore, less than 5 per cent and hardly likely to be of predictive value.

References

ELLIOTT, C. D. (1972). 'Personality factors and scholastic attainment', *Brit. J. Educ. Psychol.*, **42**, 23–32.

ENTWISTLE, N. J. (1972). 'Personality and academic attainment', *Brit. J. Educ. Psychol.*, **42**, 137–51.

EYSENCK, H. J. and COOKSON, D. (1969a). 'Personality in primary school children. 1. Ability and achievement', *Brit. J. Educ. Psychol.*, **39**, 109–22.

EYSENCK, H. J. and COOKSON, D. (1969b). 'Personality in primary school children. 2. Teachers' ratings', *Brit. J. Educ. Psychol.*, **40**, 117–31.

EYSENCK, S. B. G. (1965). *The Junior Eysenck Personality Inventory*. London: University of London Press.

RAVEN, J. C. (1965). *The Coloured Progressive Matrices*. London: H. K. Lewis.

The Reliability of Teachers' Judgement

of the Abilities Being Tested by Multiple Choice Items

R. W. Fairbrother, Centre for Science Education, Chelsea College

Summary

Twenty-two teachers were asked to give their opinion as to the ability being tested by each item in two multiple-choice examination papers. Agreement appeared to be rather low, particularly from one of the papers, but a rigid statistical analysis was impossible because of the difficulty of deciding upon acceptable criteria. The opinions of the teachers were correlated with the facility values for the items and the coefficients obtained appear to support the Bloom taxonomy.

Introduction

For some years now, the GCE and CSE Examination Boards have attempted to increase the validity of their examinations by paying closer attention to the abilities which individual questions are attempting to assess. Various publications have had an influence on this move and probably the greatest has been that of Bloom, *et al.* (1956). Bloom and his colleagues postulated a taxonomy of educational objectives and put forward their ideas in two books, one dealing with cognitive abilities and the other with affective abilities. For various reasons there has been a greater concentration of the assessment of cognitive abilities and this research looks at one aspect of this.

The Bloom taxonomy of cognitive objectives proposes a hierarchy of abilities which, in ascending order, are Knowledge, Comprehension, Application, Analysis, Synthesis, Evaluation. For further details readers should refer to Bloom *et al.* (1956). With this hierarchy as a guide, analysis of examination papers in use less than ten years ago indicated a heavy weighting towards the lower abilities with perhaps as much as 70 per cent of the marks being awarded for the ability to remember formulae, definitions, standard descriptions, etc. Only a minority of questions attempted to find out whether a candidate understood what he was doing. (See, for example, Spurgin (1967) for an analysis of advanced level physics papers.)

The present trend is towards the construction of examinations which include more of the higher abilities. Many Boards, both GCE and CSE, publish a 'declaration of intent' at the start of their syllabuses. An illustration of such a declaration is provided by the Joint Matriculation Board (1970) using four abilities for the written examination weighted as follows:

Knowledge	40%
Comprehension	30%
Application	20%
Evaluation and Investigation	10%

Proper attention is drawn to the difficulty of fulfilling the above specification accurately and 'the figures are to be regarded as giving a general indication of the intended weighting . . . there may be difficulties in assigning individual questions

---

Acknowledgements

I am indebted to the teachers who took part in this investigation. It represented an additional chore in an already overloaded day. I am also grateful to Margaret Brown and Anthony Malpas, both at the Centre for Science Education, for many fruitful discussions concerning this investigation which is similar to some work which they are doing in mathematics. Dr Robert Wood of the University of London School Examinations Department let me know about the point-multiserial correlation and did the calculations for the coefficients. This relieved me of the tedium of having to write my own computer program and I am grateful for his help and advice. The Oxford and Cambridge Board were very helpful in giving me access to data concerning the two papers.

Finally, I should emphasize that the opinions expressed in this paper are my own and do not necessarily represent those of any examination board with which I am connected.

to specific abilities and . . . questions may frequently overlap the stated categories of abilities'.

This research attempts to find out whether there is any agreement among teachers about the abilities which are being tested by an examination paper.

It should be emphasized that many Boards are beginning to give specifications similar to the above (see Schools Council, 1973) and that the JMB information is quoted because it discusses the rationale behind such a specification and because the information is conveniently available in one of the many useful papers published by this particular Board.

The investigation

For this analysis two papers were used—the 1970 and 1971 Coded Answer (multiple choice) papers of the Nuffield Advanced Physics examination available from the Oxford and Cambridge Board. Each paper contains 40 items which are intended to be answered by candidates in one hour.

Because these papers are designed for a physics course whose philosophy and aims are set out in more detail than any other A-level physics course, one might expect that if any consensus of agreement is possible at all it should happen with these papers. Yet another factor which would contribute to agreement being reached is that most of the teachers whose opinions were sought had attended briefing conferences of one kind or another which were designed to give some introduction to the content and aims of the Course. A degree of homogeneity thus seemed to be available which could not be obtained using teachers of more traditional courses.

A total of 63 teachers were contacted by letter outlining what would be involved and asking them if they would agree to take part. Thirty-eight replies to this initial letter were received and eventually opinions were collected from 22 teachers together with information on how long each teacher had been teaching the Course and how many candidates they had entered in each of the three years in which the examination had been held, up to 1972. Table 1 summarizes the years of experience of teaching the course of this sample of teachers.

Table 1: Showing the years of experience of teaching the Nuffield Advanced Physics Course of this sample of teachers

Years of Experience y	$0<y\leqslant 1$	$1<y\leqslant 2$	$2<y\leqslant 3$	$3<y\leqslant 4$	$4<y$
Number of Teachers $N=22$	1	4	7	9	1

Each teacher who took part was issued with a description and illustrative examples of four abilities based on the Bloom taxonomy. The abilities were Knowledge, Comprehension, Application, Analysis/Evaluation. The detailed description which was sent to the teachers is shown in the Appendix. A covering letter asked the teachers to get the abilities clear in their mind before working through each item and assigning it to an ability category. It was suggested that it would probably be necessary to refer continually to the descriptions of each ability before making a decision. The replies from the teachers are summarized in Table 2. Because of the small numbers involved, no attempt was made to group responses according to years of teaching the Course.

Interpreting the results

It is difficult to analyse the results in a rigid manner because there appears to be no way of arriving at a meaningful 'coefficient of agreement' between teachers for each item. If an appeal could be made to some 'infallible' judge then a start could be made by measuring how many teachers agreed with the judge. However, in this case each teacher is a judge in his own right and we have no reason to accept the opinion of one rather than of another.

There is a further difficulty in that we do not know whether the differences between adjacent abilities are all the same. If we can accept the hierarchical nature of the abilities as propounded by Bloom *et al.*, then we can postulate that pupils are more able to step from one ability to the next as they grow older. This would imply that the step from Comprehension to Application is bigger for younger pupils than for older ones. Indeed some of the steps may be too big at certain stages of development and they may never become equal in size for a fully developed individual.

It is tempting to define complete *disagreement* as being the situation which gives a question an equal number of votes for each category—a result which would be obtained if votes were assigned randomly. If there were 20 teachers, the distribution would thus be:

K	C	Ap	An
5	5	5	5

However, because an even spacing of the abilities cannot be assumed, an even distribution of votes might represent some measure of agreement rather than complete disagreement.

Remembering this limitation, however, we can assume that a rectangular distribution represents only chance agreement and apply a Kolmogorov-Smirnov one sample test (Guilford, 1965, p. 260) to the null-hypothesis that there is no agreement between teachers which is significantly better

Table 2: Number of opinions for each ability from 22 teachers

	1970					*1971*			
Item No.	*K*	*C*	*Ap*	*An*	*Item No.*	*K*	*C*	*Ap*	*An*
1	22				1	19			3
2	22				2		11	11	
3	22				3	8	5	6	3
4	17	5			4	1	13	8	
5	15	6		1	5	6	7	8	1
6	17	5			6	4	17	1	
7	16	5	1		7		6	16	
8	18	3	1		8	2	13	7	
9	16	4	2		9	10	2	8	2
10	8	14			10	15	6		1
11	12	10			11	12	6	3	1
12	8	12	2		12	7	8	6	1
13	5	9	7	1	13	1	9	10	2
14	4	9	8	1	14	5	8	9	
15	4	9	8	1	15	3	12	6	1
16	11	7	4		16	5	10	6	1
17	11	6	5		17	1	14	6	1
18	12	5	5		18	4	9	9	
19	5	6	11		19	2	10	10	
20	2	12	8		20	1	8	12	1
21	13	6	3		21		13	6	3
22	2	11	9		22	15	6	1	
23	11	4	5	2	23	13	6	3	
24	9	3	10		24	5	8	7	2
25	1	9	12		25	3	6	13	
26	9	9	3	1	26	3	9	10	
27	3	9	10		27	4	8	9	1
28	12	7	3		28	5	8	3	6
29	2	18	1	1	29	2	9	6	5
30	11	8	2	1	30	1	4	10	7
31	11	8	3		31	2	4	7	9
32	2	8	8	4	32	3	6	11	2
33	3	6	9	4	33	3	6	10	3
34		10	8	4	34	2	9	10	1
35		4	14	4	35	1	8	11	2
36		7	5	10	36	1	7	13	1
37	9	7	6		37	14	7	1	
38		4	12	6	38		6	13	3
39		14	6	2	39	1	8	7	6
40	4	13	2	3	40		3	13	6
No. of Responses (Total=880)	349	292	193	46		184	315	306	75
Fraction of Responses	0·397	0·332	0·219	0·052		0·209	0·358	0·348	0·085

Table 3: Item numbers for which agreement is better than chance at the 1 per cent and the 10 per cent levels of significance using a Kolmogorov-Smirnov test against a rectangular distribution

Year	*Significance Level (per cent)*	*Item Numbers*	*Total*
1970	1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 21, 28, 29, 30, 31	17
	10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 18, 21, 26, 28, 29, 30, 31, 35, 38, 40	24
1971	1	1, 6, 10, 22, 23, 37, 40	7
	10	1, 6, 10, 11, 22, 23, 30, 37, 40	9

than chance. The test was applied to each item and Table 3 shows the numbers of the items for which the null-hypothesis can be rejected at the 1 per cent and 10 per cent levels of significance.

It can be seen that in the 1970 paper less than half the items satisfy the 1 per cent criterion (each paper contains 40 items) and only just over half satisfy the 10 per cent criterion. Agreement is much worse on the 1971 papers.

Two further analyses were done relating the opinions for one item to the mean overall pattern of responses as shown at the foot of Table 2 instead of to the rectangular distribution. One analysis used the Kolmogorov-Smirnov test as just described, the other used Cohen's kappa (Cohen, 1960) as generalised by Fleiss (1971). The generalized kappa measures nominal scale agreement between several raters. It applies when all disagreements may be considered equally serious, and while this may not be so with this study we do not know the relative weightings to apply. In addition, because the ability categories used in this study are supposed to be hierarchical, the data are probably ordinal and not nominal. Table 4 summarizes the results of these two analyses.

It can be seen that Cohen's kappa is generally more lenient in that it accepts a larger number of satisfactory items than does the Kolmogorov-Smirnov test. The exception is in the 10 per cent level for the 1970 paper. The 1971 paper shows a considerable improvement when using kappa but the number of acceptable items is still less than half the total number.

Possible reasons for disagreement

It has not been found possible to produce a satisfactory measure of agreement between teachers; first because, there is doubt about the relationship of one ability category to the next; secondly, because there is doubt about the way in which the teachers distributed their responses; and finally, because there appears to be no statistical measure which applies to this situation.

Simply looking at the results for individual items in Table 2 indicates there is a fair amount of disagreement and it will help to consider how this might arise. One way could be caused by the teachers interpreting the instructions differently and hence basing their analysis on different criteria. It would have been better to gather the teachers together to discuss the interpretation of the instructions before asking them to analyse the questions. The question analysis itself must, however, be done independently. Festinger and Thibaut (1951) have revealed that there seems to be some pressure in group situations for members of the group to move towards agreement. Of even more relevance, Asch (1956), and Krech, Crutchfield and Ballachey (1962) report many occasions when an individual changes his mind to agree with the rest of the group even though the individual's original judgement was correct. The present research was aimed at finding out the amount of agreement between individuals working independently but basing their opinions on a common set of criteria.

Another possible cause of disagreement is that individuals solve questions differently and a problem which involves the Knowledge ability for one person involves, say, Comprehension for another. This is almost certainly true to some degree but, if its extent is as wide as is indicated by the above results, then it calls into question the value of publishing examination specifications in the detail mentioned earlier. It also raises the question as to which student we have in mind when we talk about testing for an ability and, probably more important, who makes the decisions ?

The validity of the Bloom Taxonomy

Yet another cause for disagreement calls into question the whole basis of the Bloom taxonomy. It is a classification of student behaviour and as such is independent of the subject matter which is being studied.

'... this taxonomy is designed to be a classification of the student behaviours which represent the intended outcomes of the educational process. It is assumed that essentially the same classes of behaviour may be observed in the usual range of subject-matter content as different levels of education (elementary, higher-school, college), and in different schools. Thus a single set of classifications should be applicable in all these instances' (Bloom, 1956, p. 12).

'This means that the kinds of behavioural changes emphasized by *any* institution, educational unit, or educational philosophy can be represented in the classification' (ibid. p. 14).

If an examination is designed to assess abilities in students then the Bloom taxonomy, if valid, should be capable of being used to analyse the examination, however composed and in whatever subject. If any particular ability has not been developed then students will not be able to answer those questions which test that ability. If an examination omits to test for any particular ability, whether by accident or by design, then using the Bloom taxonomy should reveal this omission.

The Nuffield Advanced Physics examinations are not composed using the Bloom categories as one of its dimensions, nevertheless if the taxonomy is valid it should be reasonable to superimpose the categories on the papers even in the form used in this research which grouped the last three categories together. Indeed, if the taxonomy is not applicable quite generally, then to make use of it to compose examinations is to impose a structure

Table 4: Item numbers for which agreement is better than chance at the 1 per cent and the 10 per cent levels of significance using a Kolmogorov-Smirnov test (KS) and also Cohen's kappa (Kappa), both tested against the mean obtained distribution

Year	*Significance Level (per cent)*		*Item Numbers*	*Total*
1970	1	KS	1, 2, 3, 4, 6, 8, 25, 34, 35, 36, 39	11
		Kappa	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 29, 35, 39	14
	10	KS	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 20, 22, 25, 27, 29, 32, 33, 34, 35, 36, 38, 39	23
		Kappa	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 20, 21, 22, 29, 35, 39, 40	19
1971	1	KS	1, 6, 10, 22, 23, 37, 40	7
		Kappa	1, 2, 4, 6, 7, 8, 10, 17, 21, 22, 23, 25, 36, 37, 38, 40	16
	10	KS	1, 6, 7, 10, 11, 22, 23, 30, 31, 37, 38, 40	12
		Kappa	1, 2, 4, 6, 7, 8, 10, 17, 19, 20, 21, 22, 23, 25, 36, 37, 38, 40	18

which is false, and is to influence teaching in the wrong direction. Such a situation could arise if the taxonomy failed to embrace an ability which was thought to be desirable, because then the use of the taxonomy to structure an examination would result in that ability not being assessed; and the use of the taxonomy to analyse an existing examination which assessed that ability would result in uncertainty as to how to categorize the relevant questions. So one interpretation of the spread of results obtained is that the Bloom taxonomy *cannot* be applied in retrospect and we must therefore question its validity.

The hierarchy of abilities

One aspect of the validity concerns the nature of the hierarchy of abilities. Such a hierarchy implies the following:

Ability Category	*Abilities Involved*
K	K
C	K+C
Ap.	K+C+Ap.
An.	K+C+Ap.+An.

In other words the higher abilities involve the lower abilities. One could pursue this by analysing individual questions and asking if, for example, a Comprehension item is spoilt by *giving* the knowledge which has to be applied instead of asking the candidate to recall the knowledge. This is not taken further in the present paper. Instead we follow Kropp and Stoker (1966) and take up the comments made by Bloom, *et al.* concerning the relationship between the hierarchy and the facility of an item:

> '. . . so long as the simpler behaviour may be viewed as components of the more complex behaviours, we can view the educational process as one of building on the simpler behaviour. . . .'
>
> 'If this is the real order from simple to complex, it should be related to an order of difficulty . . .'
>
> 'We . . . have found some evidence to support this hypothesis.
>
> Thus problems requiring knowledge of specific facts are generally answered correctly more frequently than problems requiring a knowledge of the universals and abstractions in a field. . . . 'Our evidence on this is not entirely satisfactory, but there is an unmistakable trend pointing towards a hierarchy of classes of behaviour which is in accordance with our present tentative classification of these behaviours' (Bloom, 1956, pp. 16-19).

The relationship between ability and difficulty can be investigated by finding the correlation between the ability categories and the facility values for the items in each test. If there were two ability categories the familiar point-biserial correlation coefficient could be used. As there are more than two categories a generalization of the biserial, the point-multiserial, has to be used (Das Gupta, 1960). The method involves listing all the responses of the teachers to all the items (putting K=1, C=2, Ap=3, An=4) and then setting alongside these responses the facility of the item to which they refer. This gives, two sets of 880 numbers which can then be correlated. Table 5 shows the values obtained.[1]

Table 5: Point-multiserial correlation coefficients between ability and facility. (The values are negative indicating the higher the ability the lower the facility.)

1970	1971
−0·35	−0·26

The value of the point-multiserial coefficient for each year is highly significant ($p<0{\cdot}001$) which appears to support the Bloom claim:

> 'There is an unmistakable trend pointing towards a hierarchy of classes of behaviour which is in accordance with our present tentative classification of these behaviours.'

On reflection this conclusion causes a certain amount of unease. It is obvious that one can compose very difficult Knowledge items. For example, in A-level physics: 'What is the specific heat capacity at constant pressure of hydrogen?'. Or, in Current Affairs: 'What is the name of the present head of state of Chile?'. Is the greater difficulty of items measuring the higher 'Bloomian' abilities a reflection of the intrinsically greater difficulty of *answering* such questions, or is it a reflection of the difficulty of *composing* easy questions which test the higher abilities? We can go back in time and ponder a different question which is, 'Why were examinations of twenty years ago so heavily biased towards recall?'. I suspect that, basically, it was because it is easier to ask (and mark) that kind of question than one which attempts to test other (higher?) abilities.

---

[1] The facility values are not listed because this information is confidential. I am grateful to the Oxford and Cambridge Board for giving me access to the values for the purpose of the calculations.

Implications

The outcome of most educational research is not clear cut, and implications depend upon interpretation. This particular research is no exception. It needs repeating with a larger group of teachers who have been better briefed. We should recognize also that we have asked the opinion of the *teachers* about what abilities they think their *students* use in answering the items. What do the students think? (This research is being taken further to try to answer this question).

Earlier in this paper we questioned the value of publishing detailed examination specifications if teachers disagree about the category of individual items. We also drew attention to the harm which an invalid taxonomy could cause. At the same time, however, we should consider the alternative of not using some statement of objectives. Without some monitoring procedure there would be the danger of just doing that which was most easy and assessing mainly knowledge. It should be emphasized that the examining boards only say that the examination as a whole shall be divided in a particular way. Different people may agree on the overall composition of an examination but disagree on how individual components contribute to the whole. This is a hopeful note on which to end but it needs studying further.

Appendix

Analysis of Questions into Abilities

The four ability categories into which the questions are to be grouped are: Knowledge, Comprehension, Application, Analysis/Evaluation. These categories are taken from Bloom B. S. (Ed.), (1956) *Taxonomy of Educational Objectives, 1*, Longman, and the descriptions and examples given below are based on this source. The abilities form a hierarchy such that each one is higher in the demands it makes than the previous one.

Knowledge

This ability involves remembering (i.e. recognizing or recalling) ideas, material, phenomena, etc. The answer to the question is simply regurgitated without any further processing. The ability could involve remembering:

terminology, i.e. the terms and concepts of physics;
conventions and rules, e.g. symbols on graphs or diagrams, trends and sequences;
classifications and categories, e.g. the features which decide whether a substance is a metal;
criteria;
methodology e.g. methods for determining the atomic structure of a material;
principles and generalizations, e.g. Boyle's Law;
theories and structures, e.g. kinetic theory of matter.

Example

1. Which one of the following phrases about wave motion defines period?
   - (a) The maximum distance a particle is displaced from its point of rest.
   - (b) The length of time required for a particle to make a complete vibration.
   - (c) The number of complete vibrations per second.
   - (d) The time rate of change of distance in a given direction.

Comprehension

This ability goes a stage beyond just remembering. It requires some operation to be performed in order to arrive at a solution to a problem. The data and method of solution are given or the situation is very familiar and understandable. The ability could involve:

translation from verbal into graphical form;
translation of an abstraction by giving an example;
interpretation as evidenced by inferences, generalizations or summarizations produced from given data but going beyond mere repetition and rephrasing of given information;
extrapolation beyond the given data but making use of the data to determine implications, consequences, corollaries, effects, etc.

Example:

1. When a current is induced by the relative motion of a conductor and a magnetic field, the direction of the induced current is such as to establish a magnetic field that opposes the motion. This principle is illustrated by:

(a) a magnet attracting a ball;
(b) an electric generator or dynamo;
(c) the motion of a compass needle;
(d) an electric doorbell.

Application

This ability goes beyond that of comprehension as described above. While a problem in the *comprehension* category requires the student to know an abstraction well enough that he can correctly demonstrate its use when *specifically* asked to do so, a problem in the *application* category requires the student to apply the appropriate abstraction without having to be prompted as to which abstraction is correct or without having to be shown how to use it in that situation.

In other words, application requires the student to decide for himself what bit of physics is relevant to the problem and then to correctly apply the physics to find the solution to the problem. Comprehension is simpler than application because the principle required for solving the problem is given, or the problem is sufficiently familiar so as to not require this.

Examples:

1. Suppose an elevator is descending with a constant acceleration equal to 'g'. If a passenger attempts to throw a rubber ball upwards, what will be the motion of the ball with respect to the elevator? The ball will:
(a) remain fixed at the point the passenger releases it;
(b) rise to the top of the elevator and remain there;
(c) not rise at all, but will fall to the floor;
(d) rise, bounce off the ceiling, then move toward the floor at a constant speed;
(e) rise, bounce off the ceiling, then move toward the floor at an increasing speed.

Analysis/Evaluation

These are higher level abilities than that of application and can often be assessed more satisfactorily in the Comprehension passage or in the long answer paper. However, you will find coded answer items which fall into this category.

Analysis emphasizes the breakdown of the material into its constituent parts and the detection of the relationship of the parts and of the way they are organized.

Evaluation is the making of judgments about the value, for some purpose, of ideas, works, solutions, methods, material, etc.

These abilities could include:

recognizing unstated assumptions;
distinguishing facts from hypotheses;
distinguishing a conclusion from statements which support it;
detecting logical fallacies in arguments;
recognizing which facts or assumptions are essential to an argument;
applying given criteria to the judgment of a piece of work;
recognizing and weighing values involved in alternative courses of action.

Examples:

1. Galileo investigated this problem of the acceleration of falling bodies by rolling balls down very smooth planes inclined at increasing angles, since he had no means of determining very short intervals of time. From the data obtained he extrapolated for the case of free fall. Which of the following is an assumption implicit in the extrapolation?

(a) that air resistance is negligible in free fall;
(b) that objects fall with constant acceleration;
(c) that the acceleration observed with the inclined plane is the same as that involved in free fall;
(d) that the planes are frictionless;
(e) that a vertical plane and one which is nearly so have nearly the same effect on the ball.

Important

Questions testing the higher abilities may also involve lower abilities. For example, a question testing Application may also require Knowledge, e.g. remembering terminology. In such cases the questions should be rated at their highest ability and should not be split between several abilities.

References

ASCH, S. E. (1956) 'Studies of independence and conformity. A minority of one against a unanimous majority', *Psychol. Monog.*, **70**, 9 (Whole no. 416).

BLOOM, B. S., ENGELHART, M. D., FURST, E. J., HILL, W. H. and KRATHWOHL, D. R. (1956) *Taxonomy of Educational Objectives, 1 Cognitive Domain*. London: Longman.

DAS GUPTA, S. (1960) 'Point-biserial correlation coefficient and its generalization', *Psychometrika*, **25**, 4.

FESTINGER, L. and THIBAUT, J. (1951) 'Interpersonal communication in small groups', *J. Abnorm. Soc. Psychol.*, **46**, 92–9.

FLEISS, J. L. (1971) 'Measuring nominal scale agreement among raters', *Psychol. Bull.*, **76**, 5, 378–82.

GUILFORD, J. P. (1965) *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill.

JOINT MATRICULATION BOARD (1970) *Examining in Advanced Level Science Subjects of the GCE*. Manchester: Joint Matriculation Board.

KRETCH, D., CRUTCHFIELD, R. S. and BALLACHEY, E. L. (1962) *Individual in Society*. New York: McGraw-Hill.

KROPP, R. P. and STOKER, H. W. (1966) *The Construction and Validation of Tests of the Cognitive Process as Described in the Taxonomy of Educational Objectives*. Institute of Human Learning and Department of Education, Research and Teaching, Florida State University.

SCHOOLS COUNCIL (1973) *Assessment of Attainment in Sixth-Form Science: Schools Council Examination Bulletin 27*. London: Evans/Methuen Educational.

SPURGIN, C. B. (1967) 'What earns the marks', *Phys. Educ.*, **2**, 6, 306–10.

Education 3-13

A journal for Primary Education

In the October issue

Editorial: *Colin Richards*

Children in difficulties

The needs of young children. *Ron Davie*
Teaching slow learners *Ron Gulliford*
Infant children with problems. *Lesley Webb*
Physically handicapped children in primary schools. *Elizabeth Anderson*

Other Articles

Perennial symbols; teaching through life themes. *John Hull*
Team teaching; uses and abuses. *Jennifer Nias*

It may be ordered direct from the Publishers

In Britain

The Educational Manager
Collins Educational
Kirkintilloch Road
Bishopbriggs, Glasgow G64 2PW

Annual subscription £1.50 including postage

In the U.S.A. and Canada

APS Publications, Inc.
150 Fifth Avenue
New York, N.Y. 10011

Alternatives to the Traditional Sixth Form:

a Preliminary Report

Lea Orr, formerly with the Alternatives to the Traditional Sixth Form Project, NFER, and Judy Dean, Alternatives to the Traditional Sixth Form Project, NFER

Summary
Recent years have seen the emergence of new types of provision for the 16–19 age group in both the schools and further education sectors. This paper briefly outlines the various institutions now in operation and the rationale behind each type. A four-year research project aims to survey and then evaluate the alternative arrangements for this age group. The project's first phase, which concentrated on the collection of background information and the definition of evaluation criteria, is described. Some preliminary findings are presented, together with the major issues arising from meetings and questionnaires.

I. Background

(a) *Changing provision for the 16–19 age group*

If the conventional grammar school sixth form is thought of as a model of the traditional sixth form, alternatives have been developing in recent years to many of its most characteristic features. Its highly selective intake, its purely academic and highly specialized courses, its strong emphasis on examinations and university entry, its protective atmosphere, and its generally elitist character—all these are being increasingly challenged, at least in intention, by a number of trends in the educational provision for the 16–19 age range. New types of institution are emerging within both the schools and further education sectors. There is a growing movement of 'academic' 16-year-olds from schools to colleges of further education and the overlap between schools and colleges in the courses they offer is increasing. Schemes linking or co-ordinating the provision of courses for this age group are being introduced in different areas.

At present, within the maintained sector of the system there are five main types of institution offering full-time education to the 16–19 age group. Sixth forms in 'all through' *grammar* and *comprehensive schools* are the first two. These generally differ in their underlying philosophy and stated (or hidden) objectives, and in many cases also in their range of courses and entry requirements and hence in their intakes. A third type of institution is the *sixth-form college* which caters, under school regulations, only for students aged 16–19 and may offer either purely academic courses, mainly at A-level, or a wider range of pre-vocational and non-examinations courses as well as O- and A-level. There are now some 70 sixth-form colleges operating in the country and many more are planned to open in the near future. These colleges usually have a far larger number of students than do school sixth forms and thus they offer a wider choice of courses. The great majority are 'open entry' and many of them have relaxed the traditional O-level qualifications needed for enrolling on an A-level course. A fourth type of institution is the '*tertiary*' *college*, which is a comprehensive college for the 16+, amalgamating all the sixth-form and further education provision in a given area into one college which is run under further education regulations. There are at present about 10 such colleges in the country and a few more are definitely planned. These offer the whole range of academic and vocational courses normally available in sixth forms or in further education colleges, and all are 'open entry'. The last type is the *college of further education* (other than 'tertiary'). Apart from their conventional provision of technical and vocational education, further education colleges are now often providing full-time academic courses, including A-level, for increasing numbers of young students.

(b) *Need for evaluation*

Some of the movement of pupils from schools to A-level courses in further education may be regarded as the result of voluntary choice on the part of the students themselves. The emergence of the new institutions, however, is closely associated with the reorganization of secondary education along comprehensive lines. At the local authority level, they represent attempts to rationalize resources—of money, buildings and teachers—and to accommodate viable sixth forms in the emerging comprehensive system while avoiding unmanageable school size.

These economic and administrative considerations figure, of course, in the public debate about institutional arrangements for the 16+. But they are strongly supported by educational, psychological and social arguments which concentrate on the advantages or disadvantages of alternative systems to the students themselves, and on their implications for younger pupils and their teachers. It has been suggested that the wider choice of courses available in colleges and the easy access of academic students to related vocational courses should benefit the students educationally. Other characteristics of the colleges such as the transition of students at 16, the less rigid selection at this age, the more adult environment and the social mix of academic and vocationally based students should benefit them psychologically and socially. It is also argued that the existence of the colleges is likely to encourage staying on beyond 16. On the other hand, advocates of 11–18 schools claim that the lack of continuity at 16 may discourage staying on, and that the resulting 11–16 schools may find it difficult to recruit good teachers, that college students are deprived of beneficial contact with younger pupils and of better pastoral care, guidance and counselling, which are essential at this age. It is argued that the wide age-range in 11–18 schools has a two-way beneficial effect. Older pupils develop a sense of responsibility and younger pupils have before them the example of the sixth form and a goal which is within their reach. A strong sense of belonging to a community is thereby engendered.

These are only a few of the arguments brought forward. The discussion is usually based, however, on statements of faith with hardly any evidence being offered to support the propositions made. The evaluation of the effectiveness and relevance of the alternatives should help the existing institutions and also the many LEAs which are still considering their plans for reorganization. Both planners and practising educationists are looking for evidence and for lessons that may be learned from the experience of others. Such evidence is also likely to sharpen the definition of objectives and so raise the level of argument among the many interested bodies.

II. The research project

The NFER is sponsoring a four-year project, which aims to survey, and then evaluate the alternative institutional arrangements for the 16–19 age group. Maintained institutions in both the schools and further education sectors are included in the study. The first phase of the project, carried out during 1974, was largely concerned with collecting background information on the developments in the educational provision, identifying areas of concern or uncertainty, and defining criteria for the evaluation of the emerging types of institution. In what follows, the project's activities so far and some initial findings are summarized.

(a) *Framework for evaluation—an outline*

The task of selecting criteria for the evaluation of alternative institutional arrangements has engaged the attention of the project from the outset. It has been concluded that a straightforward comparative assessment of 'output' is impracticable for a number of reasons. The ideals of at least part of the different systems, and therefore their objectives, differ. Academic attainment is stressed by grammar schools and by many sixth forms in comprehensive schools, while other comprehensive schools and many of the new colleges emphasize objectives such as a wider range of opportunities for students of a wide range of abilities and interests, vocational training, or social integration. The new colleges are few and evolving whereas many school sixth forms are fully developed and may fall back on tradition.

As far as the new colleges are concerned, it seems that their evaluation must take the form of a critical review of developing trends. The model to be adopted for the evaluation exercise is that of assessing the inputs, outputs, and the educational 'treatments' with reference to institutional types.

This would include, apart from measures of students' attainment, also an assessment of students' and teachers' attitudes and reactions to their institutions, the image of these institutions in the eyes of potential customers, that is fifth formers who are about to decide whether, and where, to pursue their education and the effect of the existence of 16–19 colleges on the resulting 11–16 schools.

The following areas of inquiry have been outlined:

Who comes into an institution and why (who does not and why)?

What does the institution offer?

How does it function?

What is its underlying philosophy and how is it being implemented?

How popular is the institution with its:
students
teachers
potential 'customers'
the rest of the education system?

What is the level of students' attainment?

To what extent are students satisfied with their course, both while taking it and in retrospect?

(b) *Sources of information*

Combining survey techniques, observations, and discussions with people involved, at different levels, in 16–19 education, the research activities during this first phase of the project can be grouped under three main headings:

(i) A survey of all Local Education Authorities.

(ii) A series of meetings with students and educationists.

(iii) A survey of some 5,000 students in their first year of post-compulsory education.

(i) *Survey of Local Education Authorities*

The main aims of this survey were to produce a picture of the current educational provision for the 16–19 age group and to identify the trends in reorganisation planned for the near future.

A questionnaire was sent to each LEA in England and Wales in October 1973 in an attempt to gather complete information before the picture was clouded by the Local Government reorganization. The questionnaire, addressed to Chief Education Officers, requested information about the size of the age group under consideration living and being educated within the local authority area, types and numbers of schools and colleges providing courses for the 16–19 age group, types of courses offered, entry requirements, future plans and general organization. A preliminary analysis, based on 114 replies (70 per cent of all LEAs) provided the findings summarized below.

(ii) *The series of meetings*

In order to identify issues of concern to those who are directly involved in 16–19 education and to define criteria by which the various types of institution might be evaluated, discussions were held with a wide range of individuals concerned with the education of this age group.

The first of these was a one-day workshop held at the NFER, to which were invited school heads, college principals, inspectors, local authority representatives, chief education officers and researchers working in allied fields.

Various possible criteria of successful education for the 16–19 age group were discussed at length: examination successes, absenteeism, students' subsequent achievement in higher education and employment, their own assessment of their education, and, at an institutional level, the degree of flexibility in subject choice, possibility of transfer between courses, pastoral care and staff attitudes. Although the participants differed in their attitudes to many of the issues involved, the points that arose suggested that the evaluation might be on three quite separate levels. These were that the study could try to assess the benefits, both academic and social, derived by students from various educational arrangements; it could review the methods of organization and operation of particular institutions and at the same time consider the total educational provision made by LEAs for the 16–19 age group.

To complement the views expressed at this workshop, the project also arranged meetings with sixth-form and college students. In each of these, a central site was chosen and representatives from local schools and colleges were invited to spend an evening exchanging views on their experiences and on 16–19 education in general. These discussions, which were left as unstructured as possible, revealed quite clearly the major educational issues that concern this age-group and also provided an insight into the characteristics of different types of institution and their effect on the students. The participants seemed to be largely concerned with the environmental and disciplinary aspects of their education and to a much lesser extent with curricular issues. A sample of the opinions expressed is presented below.

(iii) *The student survey*

A longitudinal study of a sample of students in alternative types of institutions forms a major part of the investigation. The first phase in this follow up study, carried out during the autumn of 1974, concentrated on the students' transition to post-compulsory education.

A national sample was selected of approximately 5,000 students at the beginning of their first year of post-compulsory education. They are in 45 institutions of the following types:

- Grammar schools
- Comprehensive schools
- Sixth-form colleges
- Tertiary colleges
- Colleges of further education

Each of these institutions was visited and a questionnaire administered to the students in the sample. This questionnaire asked students for details of their present course, their experience of transition to the sixth form or college, the choices they had made and their reasons for making them, their perceptions of 16+ education, their future education and career plans, and social and home background.

These students will be surveyed again at the end of their course, and much of the information already collected will be used to compare student expectations and plans with their subsequent experiences. It is also planned to follow these students through their first year in higher education or employment and examine their retrospective attitude to their 'sixth-form' course.

The head or principle of each institution was also asked to complete a questionnaire on the provision of physical, social and academic facilities, the organization of courses and extra-curricular activities, the variety of courses offered and guidance arrangements. As the main sample includes a relatively small number of institutions of each type, this questionnaire was also sent to the heads and principals of a larger sample of schools and colleges. A broader basis has thus been provided for the study of institutional characteristics.

(c) *Initial findings*

A comprehensive analysis of the data collected during the first phase of the project is in progress. The following sections present a summary of some preliminary findings from the local education authorities' questionnaires and from those completed by the sample's heads of institutions. The main issues arising from discussions with students and from unsolicited comments made by the sampled students on their questionnaires are also presented.

(i) *Students in different institutions*

The preliminary analysis of some 70 per cent of all LEAs indicated that at the beginning of 1974 the percentage distribution of full-time students aged 16–19 in the various types of maintained institutions was as follows:

Grammar school sixth forms	21
Comprehensive school sixth forms	32
Sixth form or tertiary colleges	4
Further education colleges—taking A-level	6
Further education colleges—other courses	31
Others	6
	100%

(based on 243,000 students)

As many new sixth-form and tertiary colleges have opened since these data were collected, this distribution should only be considered as a rough guide to the proportions of full-time students of this age range in different institutions.

(ii) *Entry requirements*

The information provided by the LEAs with regard to the qualifications needed for entry to the various 16–19 institutions maintained by them was necessarily of a general nature. It indicated that at the institutional level, some two-thirds of comprehensive school sixth forms are regarded as 'open entry' whereas the rest require a specified number of O-level passes for admission. All but two of the sixth-form colleges covered are defined as 'open entry'.

The relationships between the range of courses offered by an institution and its entry requirements, and the necessary distinction between the requirements for admission to an institution and those for enrolment on a particular course, especially at A-level, have been further investigated through the structured questionnaire to headteachers and principals.

Sixth forms in the majority of the sample's comprehensive schools were 'open entry'. Most of the grammar schools required a minimum number of O-level passes. For enrolment on an A-level course, the most common requirement was recommendation by the subject teacher but this is probably not independent of performance at O-level. Only about one-third of the schools demanded a specified number of O-level passes, some of them in addition to the teacher's recommendation. Of the sixth-form colleges, only two stipulated a number of O-level passes for entry to the college, but the majority required a minimum number for enrolment on an A-level course. This was also true of the colleges of further education, some of which also required certain O-level pass grades. All the colleges of further education interviewed applicants for places on A-level courses.

(iii) *A-level in colleges of further education*

As already shown, over one-third of the 16–19-year-olds in full-time education attended colleges of further education and the proportion following A-level courses in these colleges is estimated at approximately six per cent.

Half the LEAs reported an increase in the numbers of students entering A-level courses in these colleges in the last three years. Few authorities had seen a decrease in these numbers. They tended to be among those in which sixth-form colleges were recently introduced.

The majority of colleges of further education principals confirmed they had seen an increase in their numbers of A-level students over the past three years. They considered the greater choice of subjects offered in their colleges, particularly in the field of social studies, to be the main reason for this. Sociology, psychology, law, economics, and government were among the subjects most affected by this increase.

Principals were also asked the origin of students on full-time A-level courses in their college. In most colleges, about 80 per cent came direct from school, another 10 per cent had spent some time in employment and the remaining 10 per cent were overseas students.

The authorities were asked to specify any regulations (other than academic requirements) operating in their area to control entry to A-level courses in colleges of further education. The great majority stated that no such regulations existed. Apart from the few cases in which colleges required recommendations from previous schools, five authorities indicated that students were actively discouraged from taking A-level at these colleges and that 'a very good reason—medical or psychological—is needed for a pupil to transfer to college if he could go to secondary school'.

(iv) *Co-ordination of the provision for the 16–19 age group*

Both the authorities and the heads of institutions were asked whether there was a central body or a guidance committee co-ordinating the provision of courses by different institutions in their areas, and if so what were its role and composition.

About three-quarters of the authorities stated that there was no such co-ordinating body in their area. Some mentioned the separate schools and further education sections of their education committee as an obstacle to meaningful co-ordination. Others emphasized that although they had no central body, unofficial co-ordination existed in their areas in the form of 'link' courses between schools and further education colleges.

This information was endorsed by the heads and principals. Very few said that such a body existed in their area, although, in some cases there was informal liaison between the various institutions. Where such committees were in operation, they usually consisted of heads of schools, sixth-form and further education college principals, together with LEA advisers and officers.

The frequency of occurrence and the content of link courses have also been further investigated through the heads' questionnaire. All but one of the further education colleges in the sample provided courses for school pupils. The following descriptions of the types of courses offered are typical.

> 'Mostly sampling courses giving pupils an idea of mechanical and electrical engineering, fashion, distribution, catering or commerce. Some CSE link courses.'

> 'Short one-week courses introducing fourth-year pupils to a variety of careers, and CSE link courses.'

Approximately two-thirds of the sixth-form colleges and half of the schools were involved in link courses with a local college of further education. Subjects most often mentioned in this connection were shorthand, typing, metalwork, woodwork, technical drawing, surveying, law and computer science.

(v) *Specialization and general studies*

Most students and sixth formers spend only a proportion of their time on specialized studies. A typical student taking two or three A-levels or a vocational course has several periods a week devoted to what may be termed 'general' or 'liberal' studies.

The generally agreed aim of general studies is that it should counteract the specialization of most courses, broaden students' outlook and provide them with an opportunity to develop personal interests.

It is evident from data already collected by the project, that there is an enormous variation in the range, content, and structure, of 'general studies' courses available in different institutions and in students' attitudes towards this aspect of their course.

Some examples of the type of general studies courses that institutions may offer are as follows:

(1) A common syllabus which all students are required to take, leading to an A-level examination in General Studies.

(2) A common non-examined syllabus which is compulsory for all students.

(3) A range of subject options from which students must choose a specified number.

(4) A very wide range of subjects offered, attendance at any being optional.

Some institutions where students are required to attend a general studies course, also offer a range of optional leisure studies.

Almost all the institutions in our sample required students to take some form of general studies course. However, about three-quarters of the colleges and two-thirds of the schools offered a range of subjects from which students could choose. General studies was examined, either internally or externally in about half the sample's schools and colleges.

It has been indicated that colleges are offering a wider range of options than many schools. However, it seems that students' response to general studies may be more dependent on the enthusiasm of the staff involved than on the choice offered them. From our students' meetings it is clear that in some cases, staff are not convinced of its usefulness and this results in students' apathy and absenteeism. Most students seem to believe that some counterbalance to specialization is necessary but are not always satisfied that their general studies course is of direct relevance.

The project is considering and attempting to evaluate the different arrangements for general studies.

(vi) *Students' perceptions of different systems*

During the meetings held with groups of sixth formers and college students, and also from comments made by the sample's students on their questionnaires, several distinct topics emerged which seemed to be of particular concern to them. These were predominantly the discipline structure and environmental characteristics of the different types of institution they attended.

Some quotations from the meetings and the questionnaires will serve to illustrate the sort of issue that aroused most response.

Firstly, the freedom to organize one's own work was a source of satisfaction to many students in colleges.

> 'In a sixth-form college, it seems to be a lot more free. It's up to you to do the work. Obviously, if you do miss lessons, they start checking up on you, but really the onus is on you. If you want to get on, you've got to do the work.'
>
> Sixth-form college student

> 'We get to do more or less what we want. You don't really have to go to a lesson if you don't want to—there's no register here. They advise you, though, to stay and use the library, but they don't force you. You can go out if you want.'
>
> Tertiary college student.

The relatively relaxed atmosphere of colleges was also appreciated.

> 'The atmosphere at this college is much more friendly and relaxed than my previous schools and there is a good student-teacher relationship.'
>
> Sixth-form college student.

'I like college here because it's a great building and I prefer being "mixed" whereas at school we were all girls. Also I like the non-formal and friendly atmosphere.'

Sixth-form college student.

'There is much to be said for the adult atmosphere in a college of further education. The teachers treat you as adults. The opportunities are certainly available. It is up to the individual whether or not he wants to make use of them.'

Further education student.

'I enjoy coming to college as there is a much more relaxed and friendly atmosphere than there was at school.'

Further education student.

There were some dissenting voices.

'I think with being in a college, there being less discipline, you're inclined to do less work than you would at a grammar school.'

Further education student.

'Something I value in our system, I don't believe it should be really disciplined, but we have these rules and they're not all that rigid. I don't think they do you that much harm.'

Grammar school sixth former.

'Less discipline and more freedom tends to lead to an impersonal atmosphere at college. School used to have more feeling of tradition, and then it was a great honour to represent the school's sports teams.'

Further education student.

'The main problem of the sixth-form college is that it is hard to settle in with new teachers and new students. There seems to be no spirit compared to my previous school.'

Sixth-form college student.

'I prefer college to school but we have too many rules and regulations; at times it is more like school than the 5th year at my last school. I disagree with the treatment of a student who has been absent or missing from a lesson. I disagree with compulsory integrated studies. A college should be a college, not a 'scholege'. We have chosen to continue our education and should be treated as such.'

Tertiary college student.

'I like the sixth form very much and I think that if more people knew how different it was from the fifth form, more would stay on.'

Comprehensive school sixth former.

'I consider this college as just being another school, still being treated like children and being told what to do and when. When we were told about the college, they made it out to be something better than what it is. They said we would have a lot of free periods and have plenty of time to ourselves, but this is totally wrong.'

Sixth-form college student.

Many school sixth formers envied the college students' freedom to stay at home when not attending lessons.

'I don't enjoy school. I go there solely to get my A-levels. I go to lessons because I have to go to them, which I usually find enjoyable because I'm interested in them. In my free lessons, I don't like having to stay at school. It's stupid . . .'

Grammar school sixth former.

'. . . less strict on school attendance for vacant periods. Scholars should be allowed to leave school when they do not have any lessons.'

Grammar school sixth former.

College students felt that they would be better prepared to meet the demands of university life and of employment.

'I've been in a grammar school, I've been to a tech college and I've been out to work. I know what it's like to be in all three. I believe that it would have been much easier if I'd been taught and brought up in the atmosphere of a college to go out to work than if I'd been through school.'

Further education student.

'The idea of a sixth-form college is good because it allows boys and girls of varying educational and social backgrounds to mix and exchange opinions, so giving people a wider outlook on life.'

Sixth-form college student

'If going on to higher education, then a college environment makes the transition from school to university or teacher training college easier. As most courses at college last one or two years, you have to learn to mix and make friends in a short time—this helps to develop your character.'

Further education student.

Some school sixth formers resented the assumption that they led a sheltered life and stressed that they had a wide circle of friends, both in and out of school.

'I go to school and in my outside activities, I mix with people that are ordinary and I get just as much out of life. I don't see that going to college would broaden my outlook. I meet a cross section.'

Grammar school sixth former.

The question of prefect duties in schools and responsibilities towards younger pupils aroused much comment.

'It gives you a sense of responsibility. You feel that when you go out in the world, you'll be able to cope much more easily.'

Grammar school sixth former.

'I believe the sixth form to be a good thing. We are given responsibility and a chance to lend a helping hand in the running of the school. It also gives younger members of the school a chance to look up to and follow a good example.'

Comprehensive school sixth former.

Some college students felt that the academic work produced just as much of a sense of responsibility, and expressed the view that prefect duties were 'just keeping kids out of teachers' way'.

Uniform was a sore point for many school sixth-formers.

'I think it is a great enticement—the freedom—especially in our school. Lack of school uniform would be a big step forward.'

Grammar school sixth former.

'A lot of people—the reason they come to college is not necessarily that they think college is going to be any better. They know they'll probably get the same grades in exams as at school—but they know they won't have to wear a uniform.'

Further education student.

'School uniform shouldn't be worn as it is generally very expensive and unsuitable.'

Grammar school sixth former.

Many believed that the relatively small numbers in school produced a greater community feeling.

'Surely, that's the advantage of our school. 'We've got one headmistress, one deputy head. If you've got a complaint you go straight to the head and it's all very personal.'

Grammar school sixth former.

'I find this place very big. I know it's a help in preparing you for university, but I think I'd like a sixth form more. I think a headmaster would be able to write your report and know a lot more about you. I mean, the principal, he couldn't write anything about me, or even my group tutor.'

Tertiary college student.

'The college is so large that it's hard to get to know anybody.'

Sixth-form college student.

The number of comments made about facilities and amenities in the schools and colleges indicates that this is another important issue.

'In our sixth form there is a great need for more facilities. I visited just one technical college and I found the facilities there most interesting and pleasing. The lack of facilities leads to disorganization, leading to work not getting done.'

Grammar school sixth former.

'The sixth-form college I attend offers more amenities and better social and practical opportunities than a school sixth form.'

Sixth-form college student.

'The working facilities at my grammar school were disgusting. The few rooms that were free were crowded out etc. Scientific apparatus was verging on antique and continually being broken and not replaced.'

Further education student.

'The school has a good library and study booths that are great for private study.'

Comprehensive school sixth former.

The vast majority of young people expressed loyalty to the institution they attended. Many who were vocal in their complaints stressed that they enjoyed school or college life.

'Although I still complain a lot about the work, I'm very happy at school. I will be sorry to leave.'

Grammar school sixth former.

'Even though lack of money is a great problem to students, I wouldn't change places for the world with some of my friends at work as I am enjoying myself no end here and obtaining higher qualifications than them.'

Further education student.

'Although I have said I didn't enjoy the fifth form at school, life in the sixth-form is much better and I like being made responsible and treated as an adult.'

Comprehensive school sixth former.

Finally, the issue that probably aroused more comment than any other from students in all types of institution was that of grants. The quotations that follow give some indication of the feelings expressed.

'I believe there is a good case for sixth formers to be entitled to some "pocket money" in the form of a small grant. This money at present has to be earned on Saturdays, effectively making a six day working week regardless of homework. Of course, the average sixth former will be supplied by his parents but they have the cost of supporting him anyway and the money supplied, if it can cover necessities, does not go far towards promoting a social life.'

Grammar school sixth former.

'A grant is needed by sixth-form pupils. This would enable them to have some independence and a small social life. Contact is lost between pupils and their working friends through lack of funds.'

Comprehensive school sixth former.

'I think that the educational grant for pupils should be raised, and with the result many more pupils would be encouraged to stay on at school and gain more qualifications.'

Tertiary college student.

'There should be some type of grant given to students in sixth forms similar to the grants of students at universities and colleges of education.'

Sixth-form college student.

'I disagree with the grant of £15. Our equipment costs £17, then £3 for overalls plus college bus fares and dinner money. I think the grant should cover it.'

Further education student.

In addition to the views quoted above, some interesting indications of students' attitudes appeared during piloting of the students' questionnaire.

When asked whether they considered it preferable for people over the age of 16 to be educated in separate colleges or in schools, and for what reasons, the responses echoed the views already expressed. Those in favour of schools largely said they thought it preferable to be in a familiar environment, that schools provide training in leadership and that younger pupils benefit from seeing what sixth form is like. The most frequently mentioned arguments for colleges were that they provide a more adult environment and a more suitable atmosphere in which to study.

College students were asked whether there had been a sixth form at the school they attended and if so, why they left. Some gave the reason of inability to take the subjects they wanted at school, but the majority indicated a general dislike of school or the wish to study in a more adult environment.

Students were also asked whether they considered it desirable for people to take a job for a year or so at the age of 16 and then return to full-time study. The majority were not in favour of this, saying that they thought people would be less likely to return to full-time study, having got used to being financially better off and out of the habit of studying. Those in favour thought that work would give a greater appreciation of the benefits of education.

Many of the issues outlined above are being further investigated by means of both structured questionnaires and discussions with students and teachers.

(vii) *11–16 schools*

It has become clear during discussions that the arrangements made for educating the 16–19 age group have direct implications for those pupils still in compulsory education. Three separate issues have been raised in discussions. The first is the place and value of a prefect system in which older pupils assume some responsibilities for the conduct of the younger; the second is the importance or otherwise of the continuity that an 11–18 school provides in offering advice and guidance to the individual student and in the example that the sixth form sets for younger pupils. The view has been expressed that the absence of a sixth form may have a detrimental effect on staying on. The third is the question of teacher recruitment for the 11–16 schools. It has been suggested that without the attractions of sixth-form teaching these schools will inevitably recruit teaching staff of a lower standard.

Therefore, attitudes of fifth-form pupils and their teachers to alternative systems at 16+ are also being examined.

Further Education

an Alternative Route Reviewed

G. M. Young, Department of Adult Education, University of Manchester, and M. F. McDonald, Department of Commerce and Economics, Wigan College of Technology.

Summary

This paper reports a small scale research at a single technical college. It places the research against the background of opinions which see the colleges of further education as providing an alternative and more egalitarian route to professional employment for the failures of the selective system of secondary education. This view argues for the existence of a number of alternative routes to white collar employment through the colleges of further education. As part of an empirical examination of the existence of these hypothesized routes and their sociological significance, the authors undertake an analysis of the relationship between social and educational background and achievement among 101 students taking full-time GCE Advanced level Social Science and Arts courses at a South Lancashire technical college. The general finding, within the limited framework of the research, was that, although the college in question appeared to be offering to some students a temporary alternative route to higher education and professional employment, there seemed little evidence that this route was characterized by remedial or egalitarian features of any consequence.

1. The alternative route

The Crowther Report (CACE, 1959) states at one point that '. . . one of the strongest points of the further education system is that it never closes the door or refuses a second chance to anyone who has the persistence to continue or the ability to succeed'.

The notion that colleges of further education provide a passage to graduate and professional status for children, particularly of manual workers, who have not had the benefit of selective education —what has come to be known as the 'alternative route' conception of further education—is a recurring theme in the debate about the role of such colleges in the educational system as a whole. Recently, however, the conception of the alternative route has been subjected to critical attack. Unearthing data from a number of sources, Hordley and Lee (1970) have found a disproportionate number of students from selective schools on the more advanced courses and of students from middle class backgrounds on the full-time courses and those which prepare students for professional occupations. The conclusions which they have drawn from the evidence available would seem to offer little encouragement to apologists for the 'alternative route' theory. 'Substantial under-achievement among early leavers and non-selective pupils can' they suggest, 'only be eliminated by an educational system which, somewhere, has a much more powerful "remedial" orientation'.

While not wishing to dissociate ourselves from the conclusions which Hordley and Lee draw from a very thorough review of the problem, we feel the time is ripe for a fresh look at what we mean by the term 'the alternative route'. We need to ask first of all what the route is to which we are referring, and, secondly, to what is it an alternative? The first question we have already answered. The route in question is the route to graduate, graduate-equivalent and professional occupational status which lies outside the main system of selective education. The second question is more difficult to answer and in attempting to answer it we are led to look much more closely at the relationship between the colleges of further education and the selective system.

The main route to graduate and professional status lies through a process of gradual educational selection in which at certain points large numbers of individuals either withdraw voluntarily or are eliminated. The critical stages occur at the point of access to certain institutional areas: the selective schools, the sixth forms of those schools, and the universities and polytechnics, the institutions of higher education. Now, if we turn to the colleges of further education we find that the courses which they offer are so wide and varied, extending from operative training to professional preparation, and the points of entry to them are so numerous, that they can provide an institutional alternative to any or all of the phases of selective education. In fact, by using as a criterion the different points of entry to and exit from the further education system, we can delimit not one but a number of possible alternative routes. Each of these routes we can illustrate by describing the sort of academic career which would typify a successful negotiation of it. The 'success' which concerns us is the successful acquisition of formal educational qualifications. It is certification rather than attainment that has excited the interest of believers in the existence of an alternative route.

The first type of academic career is exemplified by the student who leaves a non-selective secondary school with few, if any, formal qualifications and, after taking up an industrial apprenticeship, eventually by part-time and evening study overcomes the succession of hurdles to professional status. The core of such a career has traditionally been the acquisition of an Ordinary and Higher National Certificate in one of the branches of engineering. Venables (1972) has documented the careers of a large sample of students who pursued this form of study in the late fifties and early sixties. The majority of her examples were early leavers mainly from non-selective schools, and the arduous nature of their journey through the national certificate route is indicated both by the heavy failure rate which she records and the length of time—in some cases as long as 12 years—it took for many to achieve success.

The Report of the Advisory Council on Scientific Policy (1959), which investigated the qualifications of admission to professional institutes in 1958, found evidence which suggests that many professional technologists have risen from unpromising beginnings to qualify through the national certificate route. This route we might term 'the "extended" alternative route' since through it the colleges of further education offer an alternative not only to the sixth form and the university but also, for the early leaver from a secondary modern school at least, to the selective system as a whole. It is about this extended route that Hordley and Lee are increasingly pessimistic. It is, they believe, no more than a shadow of what they and others would wish it to be.

The second type of academic career is also that of an 'early leaver' but of a different kind. It is the career not of a secondary school 'failure' but more often of a secondary 'success'. It is the career of a student who, often after achieving reasonably good 'O' level or CSE results at school, decides on the basis of immediate short-term considerations to continue his education at a college of further education for a limited period of time. He may be a student from a secondary modern school who can only find at the local 'tech' the courses he requires, or a grammar school boy who finds the climate or curricula of his sixth form too constricting, or even, as Lawrence (1973) found, a public schoolboy whose parents can no longer afford to pay extended fees. It is along this route through the colleges, we suggest, that we are likely to meet the greatest diversity in student backgrounds and interests. However, beneath the diversity runs a common ambition. A recent study at a Lancashire college (Young 1971) has confirmed what the Schools Council Sixth Form Survey (1970) earlier found, that the majority of such students entertain occupational aspirations (e.g. for teaching and the law) which can usually only be realized at institutions of higher education. Sometimes the type of student we have described will be found on industrial or commercial ONCs and ONDs—a route (although outside the sciences and technology not a very important one) to a CNAA or university degree. A more popular choice will be a GCE 'A' level course, particularly one on the Arts and Social Sciences side, the major route to institutions of higher education. It is interesting that in a period of relative stagnation and retrenchment in the colleges the number of students taking GCE 'A' level courses at grant-aided establishments of further education has steadily increased, from 52,838 in 1968 (a figure which includes the contribution of the then regional colleges) to 57,622 in 1969, 62,248 in 1970, and 67,885 in 1971 (*Statistics of Education*, 1969, 1970, 1971, 1972). This, the second of our alternative routes, we might term the 'temporary' or 'medial' alternative route. Through it the colleges of further education provide an alternative to a limited sector of the selective route, the sixth form of the selective school.

The third and last type of academic career is exemplified by that of the selective school sixth former who, after taking his Advanced levels, either fails to matriculate or even with matriculation chooses not to go to a university or polytechnic and opts for a commercial or industrial HND or HNC or professional qualification, or even, although this is becoming increasingly rare since the proliferation of the polytechnics, a more vocationally biased degree at a college of further education. As a phenomenon this type of career has undoubtedly existed for some time. Robbins (Committee on Higher Education, 1963) found that only 60 per cent of those with university entrance qualifications actually entered university, a further 24 per cent going on to teacher training or further education, the majority to the latter. The same report also records that 71 per cent of HND students had attended selective schools and 38 per cent had one or more 'A' levels. This alternative route we might call the 'final' or 'higher' alternative route in which the college of further education plays the role, often ascribed to it in folklore, of the 'poor man's university'. With the channeling of Advanced level courses away from the further and into the higher education sector we assume that this route, even more than the extended route, will gradually become a thing of the past.

Having offered a tentative classification of the roles which colleges of further education might be seen to play as channels of social mobility, we need to make a number of qualifications. First of all, the majority of young adults from non-selective schools who experience any form of further education do so as craft apprentices and semi-skilled operatives on some form of day-release. For this enormous group as Hordley and Lee have remarked, the alternative route is little more than a shadow of what it ought to be. In this paper, it seems, we are concerned with a minority of the further education student population. Secondly, in looking for operating variables in the background data of students along the route we have described, we need to make a careful distinction between the influences of social and educational backgrounds. The reason for this is that if we take the average correlation between social class and academic achievements as lying somewhere between 0·30 and 0·35 then we are forced to admit, along with Miller (1971) that nearly 90 per cent of the variance in achievement is left to be accounted for by other factors. In other words by late adolescence we can expect the influence of factors other than social class to have worked its way into our achievement data. We can expect the phenomena of the middle class 'failure' and the working class 'success' to have emerged with greater clarity than at the secondary stage. Since the apologists for the alternative route perceive the colleges as playing both an egalitarian and a remedial role, either the social class background or the prior educational achievement of incoming students can be used to strengthen or weaken their case. For example, if we were to find large numbers of students from working-class or non-selective backgrounds negotiating the extended route we have described, this would strengthen their case. On the other hand, if we were to find few of this type of student but many from middle class or selective school backgrounds passing through the 'medial' or 'final' routes, this would weaken it. A third point to be borne in mind is that an 'open door' policy with regard to the admission of students to courses does not by itself establish the status of the colleges as retrieval mechanisms for the selective system. We would also wish to know whether, after access to a college course has been obtained, there is a correlation between the background variables outlined above and the student's later academic success. Finally, it is worth remembering the heterogeneity of levels and courses in further education which Peters (1967) and others have described. The internal environments of colleges of further education probably contain a greater diversity of elements than do those of even the largest schools. It should not surprise us, therefore, if we were to find in any one college, or even, in certain cases, on any one course, different groups of students from different backgrounds pursuing more than one alternative route. It is to this question that we now turn.

2. A South Lancashire College

Between 1971 and 1973 the authors analysed the background and achievement of 101 students taking full-time GCE Advanced level Social Science and Arts courses at a South Lancashire technical college. Forty seven of the sample were men, 54 women. All but ten were between 17 and

Table 1: Parental Occupation of students on full-time GCE advanced level courses in further education (per cent)

Classification: Registrar-General's index of occupations. classification: 1966		*101 Students at a South Lancashire Technical College*	*Young (1971) social background of 'A' Level students at Burnley Technical College*	*School Conncil Sixth Form Survey, Vol. II: Further education (1970)*	*Married males in total population*
I.	(Prof.)	7%	21%	18%	4%
II.	(Man. and Sen. Tech.)	40%	30%	41%	18%
III.	Non-manual (Cler. and Sales)	6%	41%	11%	
III.	Manual (Skilled and Forem.)	33%		23%	52%
IV & V.	(Semi- and Unskilled Manual)	13%*	8%	8%	26%

*2% unclassified

20 years of age. The college studied offered a wide range of full- and part-time courses in the natural sciences, engineering, art, management, secretarial and commercial courses; approximately one-third of its full-time students were taking GCE courses at either Ordinary or Advanced level. The catchment area contained a number of sharp environmental contrasts but the bulk of the student body was drawn from the predominantly working-class areas of the South Lancashire towns. An earlier study of a neighbouring college undertaken as part of another research project had revealed high levels of occupational aspirations among students from diverse backgrounds on full-time 'A' level courses (Young, 1971). It was here, we felt, on this type of course, that we might find students from different backgrounds attempting to use the college of further education as a 'temporary' or 'medial' alternative route to graduate and professional status. In the following pages we describe the background of our students and its significance for their ultimate success.

If we turn to Table 1 we find, first of all, an over-representation of students from non-manual backgrounds in our sample. Over half of our students came from such backgrounds and, although not as pronounced as that found by the Schools Council Sixth Form Survey, which re-

Table 2: The school background of students on full-time advanced level courses in further education

Type of School	*101 students at a South Lancashire Technical College*	*Young (1971) social background of Full-Time 'A' Level students at Burnley Municipal College*	*Lawrence (1973) 'Further education students: the forgotten Sixth Form*	*Schools Council Sixth Form Survey, Vol. II: Further Education (1970)*
Independent and direct grant	10	8	28%	29%
Selective grammar	51	33	26%	27%
Secondary modern	36	39	42%	35%
Other	4	21	4%	10%
Total	101	101		

ported a higher percentage of middle-class students in the technical colleges than in the sixth forms of grammar schools, the over-representation is in line with the findings of earlier research and appears massive when seen against the background of the predominantly working-class catchment area of the college.

The notion of the open door college is again brought into question when the school background of our sample is analysed. Of the 101 students the majority, it appears had been to selective schools (either Grammar, Direct Grant or Independent), a finding which could have been predicted from any one of a number of similar researches (Young, 1971; Lawrence, 1973; Schools Council Sixth Form Survey, 1970). Initially, we were tempted to draw from the data of Tables 1 and 2 the conclusion that the college we were studying was serving largely as a 'second chance' institution for failed middle-class selective school children whose parents were striving to prevent the downward mobility of their offspring. In this we would have been mistaken. A chi-square analysis relating school attended to parental occupation revealed no significant relationship. (Value of χ^2: 1·409 with 1 df.)

As can be seen from Table 3, a majority of students from both manual and non-manual backgrounds had previously attended selective schools.

Table 3: School background related to parental Occupation for 101 full-time GCE advanced level students at a South Lancashire technical college

School attended	*Parental Occupation*		
	Manual	Non-manual	
Selective secondary	25	35	60
Non-selective secondary	21	18	39
Totals	46	53	99*

*2 unclassified by parental occupation

To gain admission to the course students were required to have four 'O' level passes on their equivalent, although exceptions were made in the case of older students. By converting pass grades at 'O' level to numerical scores (e.g. Grade 1 Pass=6 points, Grade 2 Pass=5 pts., Fail=0 pts.), we were able to analyse the relationship between the previous educational achievement of our students and the background data above. Perhaps not surprisingly, in view of what we have said earlier about the relationship between educational achievement and social background at this level, we found school background (χ^2=8·2460 with 2 df) and not parental occupation (χ^2=5940 with 2 df) to be significantly related to previous achievement at 'O' level.

By using the same procedure, i.e. converting grades into numerical scores, we were able to look at the statistics of the relationship between our background variables and the achievement of the students at the end of their course. In the end we were unable to find any relationship of significance between school background and achievement on the course (χ^2=3·32 with 2 df) although we found a strong overall correlation between the latter and previous 'O' level scores (r=+0·577, Pearson Product Moment). Of more interest to us, however, was the reappearance in the data in Table 4 of a phenomenon which had not been apparent in our sample on entry, a significant relationship between parental occupation and achievement (χ^2=12·67, p=·001, 2 df).

Table 4:

GCE 'A' level performance	*Total no. of students from manual backgrounds*	*Total no. of students from non-manual backgrounds*
0–4 pts. (failure to bare pass in 2 subjects)	30	20
5–8 pts. (a predominance of C and D grades)	6	20
8+ pts. (higher grades of pass)	10	13

The resurgence of the social class factor seemed to go against the grain of thinking on further education since Crowther, who tended to dismiss it as a factor related to achievement once access to a course had been obtained. We were moved, therefore, to delve deeper into our data in a search for possible explanations. Analysis of grades in the five leading subjects taken at 'A' level showed that performance in individual subjects was totally unrelated to social background (in fact, subject by subject, the students from manual backgrounds achieved marginally higher mean scores). Gradually it became clear to us that the key variable was related to subject choice and that those students who gained good overall grades did so partly because they chose the most favourable combination of subjects. Could subject choice, we asked ourselves, be related to parental occupation? In three of the five main subjects, economic history, sociology, and geography, there were no significant differences either in results or in the proportion of students entering from different social backgrounds. In two subjects, however, economics and English literature, a t-test revealed significant differences in performance ($t = 8{\cdot}108$, $p = {\cdot}001$, 18 df). Although performance in both subjects was unrelated to parental occupation (that is, the students from manual backgrounds who took the economics exam did just as well as their middle-class peers, and equally badly in English literature, a far greater proportion of middle-class students chose to take economics, the more favourable subject (23 out of 53). It was the bias in subject choice which accounted for the depression in the overall achievement of our working-class subjects. Nor did this bias represent the vagaries of a single year. A follow-up study of the same college two years after the original sampling produced a similar pattern in terms of a heavy middle-class bias towards economics, and similar differences in achievement. The consistency with which our working-class subjects tended to opt for the traditional arts subjects seems important in view of the tendency, noted by Lawrence, of colleges of further education to specialize in other, particularly the social science fields. It may be that students of this kind are failing to benefit from what the colleges have to offer in the way of expertise.

---

*In view of the fact that the chi-square and 'L' tests of significance are usually taken to apply to simple random samples from infinite populations the reader is advised to treat their use here as mainly of illustrative significance.

3. The alternative route reviewed

Earlier we described a number of ways in which colleges of further education might provide for certain types of students 'an alternative route' to professional occupations outside the main selective system. We also suggested that the case of those who saw the colleges playing an egalitarian and remedial role in the educational system would be either strengthened or weakened by research into the social and educational background of students on certain courses and into the relationship between background and achievement. The college and the course we studied offered us the chance to observe at first hand the workings of one of three alternative routes to higher education which we had outlined in the college environment. This was the 'temporary' or 'medial' route in which the colleges function rather like the grammar school sixth form as a bridge between the school and the institutions of higher education.

Although our original target sample was small, and representatives of only a limited sector of both the college and the further education student population, we were able to relate some of our findings to broad trends within the college over a number of student generations. Our conclusion is that, although the college course in question was offering to some students a temporary alternative route to higher education, within our limited research design we could find little support either in the statistics of access or in those of survival for the view that this route had remedial or egalitarian features of any consequence. Despite the location of the college in a mainly working-class area, the composition of the entering student boy on the 'A' level course was characterized by a predominence of students from either middle-class or selective school backgrounds. It would have been interesting to have compared it with the social class composition of the local grammar school sixth forms. In one respect our achievement data might offer some consolation to those who maintain that courses of this kind can act as a retrieval mechanism for the selective system. Where final success on the course was concerned there was no advantage in our sample in a student coming from a selective school background. However the fairly strong correlation we found between 'O' and 'A' level results suggests that there was some advantage in having reasonably good 'O' level grades on entry. 'Even more interesting, there was also an advantage, it seems, in coming from a non-manual background. The reappearance of the social class factor in achievement at this level caught us unawares. The tendency of writing on further education since Crowther had been to attempt to diminish its importance. Our suggestion was that the advantages of the student from a non-manual home stemmed from know-how and technique in choosing his combination of subjects. Certainly there were interesting social class differences in subject choice which persisted in the college over a number of years. Overall, the findings give little ground for optimism to theorists of the alternative route although we are forced to qualify them by pointing to the smallness of the sample, the numerical limitations of the ultimate population, and our own failure to control for the ability scores of our sample. After all, as Venables has said, if failure and drop-out in further education could be shown to be related to ability, the system, though punitive, could still be seen to be fair. So far, as she is quick to point out, the evidence for this is thin.

References

Advisory Council on Scientific Policy (1959). *Scientific and Engineering Manpower in Great Britain, 1959*. London: HMSO.

Central Advisory Council for Education (1959). *15–18*, Vol. I. London: HMSO (Crowther Report).

Great Britain Committee on Higher Education (1963). *Higher Education*. London: HMSO (Robbins Report).

Hordley, I. and Lee, D. J. (1970). '"The alternative route"—social change and opportunity in technical education', *Sociology*, 4, 1, 23–50.

Lawrence, P. (1973). 'The forgotten sixth form', *New Society*. *Dec*.

Miller, G. W. (1971). *Educational Opportunity and the Home*. London: Longmans.

Peters, A. J. (1967). *British Further Education: A Critical Textbook*. Oxford: Pergamon Press.

Schools Council (1970). *Sixth Form Survey, Vol. II*. Students on full-time courses in colleges of further education. London: Books for Schools Ltd.

Statistics of Education (1967, 1970, 1971 and 1972). *Further Education—England and Wales, Vol. 3*. London: DES and HMSO (for 1968, 1969, 1970 and 1971).

Venables, E. (1972). 'The human costs of part-time day release', *Higher Education*, 1, 3.

Young, M. E. (1971). 'A survey into the social and educational backgrounds, the aims and achievements of full-time "A" level students at Burnley Municipal College'. Manchester University DASTE Dissertation.

The Educational Needs of the Mentally Handicapped

John M. Hughes, Department of Special Education, Caerleon College of Education

Summary

The educational needs of mentally handicapped children in the 'new' day special schools are examined from three major perspectives. First, as a result of the examination of questionnaires received from headteachers of 88 schools; second as a result of the examination of the skills of 151 mentally handicapped children in 44 of these schools, and third as a result of an examination of questionnaires completed for 52 parents of mentally handicapped children. The results of the investigation highlight the effect of these schools being in the 'educational wilderness' until April, 1st, 1971; the low level of achievement in certain social skills, and parental problems and attitudes.

Introduction

The Education (Handicapped Children) Act, 1970 provides for the education of mentally handicapped children to be the responsibility of local education authorities, and took effect on 1st April, 1971. So, 100 years after the passing of the Education Act of 1870, universal compulsory education became a reality. The responsibility for the mentally handicapped in junior training centres passed from the Department of Health and Social Security to the Department of Education and Science. So every child who had previously been found 'unsuitable for education at school' was able to receive educational treatment.

There have been very few longitudinal studies on mentally handicapped children which have been concerned with the educational needs and development of these children. Kirk (1965) summarizes a number of American studies, while useful suggestions for a teaching programme have been made by Neale and Campbell (1963) and Russell (1964). Cain and Levine (1961), in their American study, compared 'mentally retarded' children in training centres, hospitals and at home. In the British Isles, Simpson (1967) comments on a survey and the many visits he has carried out to many junior training centres over a period of five years—'Patently, it is obvious to a trained educationist that aims, both general and specific, are conspicuous by their absence throughout large areas of the training centre world'. Marshall (1967) and Norris (1968) have produced two very important studies and provide some insight into the work being carried out in the 'new' day special schools.

Aim of the investigation

The main objective of this investigation was to ascertain the educational needs of the mentally handicapped in the 'new' day special schools (formerly junior training centres). It was hoped to attain this objective by means of investigating the facilities existing in a random selection of schools in England; the educational programmes in these schools; headteachers' opinions of their children's needs; and parental opinions. It was further hoped that an assessment of the skills of a group of children randomly selected from these schools would provide some indication of the attainments of mentally handicapped children and, at the same time, provide some indication of the needs of these children if one accepts that the chief aim of education for the mentally handicapped is to develop his individual personality and help him become as socially competent and as socially acceptable as possible.

The design of the investigation

The investigation was divided into three main sections as follows:

1. Headteachers' Questionnaire
2. A Skills-Assessment Sheet
3. Parents' Questionnaire.

1. *Headteacher's Questionnaire*

This requested information on many aspects of the school including pupils and staff; general facilities; buildings; daily activities, and headteachers' opinions of children's needs.

Selection of schools. A list of 'new' special schools in England was obtained from the Special Education Branch of the DES. The hospital schools were excluded and 40 schools were selected at random for a pilot investigation. Thirty-one (77·5%) questionnaires were returned. As a result of the pilot investigation, a few modifications were made to the questionnaire. The 40 schools were excluded from the list of special schools and a further 115 schools were selected at random as the final sample. Eighty-eight (76·5%) completed questionnaires were returned. As a result of the returns one was able to study the educational conditions and facilities for 6,058 mentally handicapped children aged from 2½ to 17 years.

2. *A skills-assessment sheet*

This requested an assessment of 60 skills of social competence.

As a result of studying three scales for assessing social skills, i.e. *The Progress Assessment Charts* (P.A.C.), Gunzburg (1963 and 1965); *The Vineland Social Maturity Scale*, Doll (1965); and *The Manchester Scales of Social Adaptation*, Lunzer (1966), the investigator decided on four main areas of social growth: Self-Help, Communication, Socialization and Occupation.

The four main areas contained the following 60 skills:

1. *Self-help*
1 Takes a drink from a glass without help
2 Eats independently
3 Has bladder control
4 Has bowel control
5 Asks to go to the toilet
6 Toilet-trained and cleans himself
7 Walks independently
8 Walks upstairs independently
9 Walks about neighbourhood independently and crosses streets (unsupervised)
10 Uses public transport independently
11 Dresses independently
12 Washes independently
13 Brushes his teeth
14 Brushes or combs his hair

2. *Communication*
15 Has a good vocabulary of isolated words
16 Carries on conversation with self or toys
17 Uses simple sentences
18 Obeys simple instructions
19 Speaks without much hesitation
20 Repeats a story without much difficulty
21 Differentiates between big/little, long/short, etc.
22 Matches primary colours
23 Names primary colours
24 Counts mechanically to ten
25 Counts five objects
26 Names coins up to 10p
27 Gives change out of 5p
28 Tells right from left on himself
29 Prints or writes own name and a few simple short words
30 Recognizes and understands 30 or more words by 'Look and Say'
31 Tells time to 'quarter past' and 'quarter to'
32 Names days of week
33 Knows own birthday (day and month)
34 Knows his age
35 Knows his parents' Christian names

3. *Socialization*
36 Plays co-operatively with others
37 Is active member of club, team or other organization
38 Uses costumes and materials
39 Plays competitive games
40 Acts out stories
41 Plays team games, e.g. football
42 Helps in domestic tasks
43 Takes on minor responsibilities
44 Carries out routine tasks by himself
45 Shows other children how to do a classroom job
46 Goes on simple errands outside the home and school
47 Makes minor purchases
48 Respects rights and property of others

4. *Occupation*
49 Uses scissors
50 Uses hammer correctly
51 Builds a tower of 5 to 7 bricks
52 Threads beads
53 Takes a nest of cubes apart—puts together again

54 Uses pencil for recognizable drawings
55 Occupies himself—does not require constant supervision
56 Uses plasticine, building bricks, etc. for construction purposes
57 Rides a bicycle in school surroundings
58 Rides a bicycle in home surroundings
59 Throws a ball and hits an object (e.g. a foot square) from about two yards
60 Uses playground apparatus in fairly safe manner

Selection of children. A pilot investigation was carried out in four local special schools. It was decided to exclude those children in 'special care units' and those who were severely physically handicapped. It was also decided to categorize the children under four headings as follows:

1. under 8 years
2. over 8 but under 11 years
3. over 11 but under 14 years
4. over 14 years.

A balance was ensured between the sexes, and two children from each school were assessed. As a result of the pilot investigation, a few modifications were made to the Skills-Assessment Sheet. Fifty schools were selected at random from the 88 schools for the final sample. Four Skills-Assessment Sheets were sent to each school—one sheet for each age category. Forty-four (88%) schools returned completed sheets and 151 children were assessed.

3. *Parents' questionnaire*

This requested information on such factors as social and medical facilities available for parents and children; parents' greatest concern; attitudes towards educational facilities.

Selection of parents. The procedure adopted for selecting parents for a pilot investigation and eventually used for the final investigation was similar to that used for the Skills-Assessment Sheets, i.e. special care children and those who were severely physically handicapped were excluded and the same age categories were used. As a result of the pilot investigation, minor modifications were made to the format of the questionnaire. Students assisted the investigator in the interviewing of 52 parents from various parts of England and Wales.

The main results

A. *The 88 new day special schools*

1. There were 6,058 pupils aged from 2½ to 17 years with 2,118 (35%) mongols. The average number of pupils per school was 68·8 and the average number per class, 11·7.

2. Three-hundred and eighty-three (66%) of the 590 teaching staff (including part-time staff) were qualified. The teacher–pupil ratio was 1:10·7 (1:8·0 when full-time teachers' helpers were included). 72 (86%) headteachers were qualified.

3. Fifty-four (61%) schools were purpose-built; 12 (14%) were in old school buildings and 22 (25%) were in converted buildings. 64 (73%) schools had adequate playground facilities and 56 (39%) had adequate storage facilities either indoors or out.

4. Only 16 (18%) schools had a gymnasium. 60 (68%) had gymnastic equipment. 56 (64%) schools had use of a swimming pool. There appeared to be an adequate supply of classroom equipment.

5. Eighty-four (95%) schools arranged outside visits and 58 (66%) invited people to the school to talk to the children.

6. Only 51 (58%) had facilities for the under-fives.

7. Sixty-one (69%) had formal timetables and 40 (66%) of these had set times for separate activities. The remaining 27 (31%) either emphasized flexibility in timetabling or indicated apparent flexibility.

8. The average number of minutes per school week for 2,782 children in 40 of the 88 schools was 1,825. 49 per cent of this time was allotted to non-teaching activities. 20 per cent of nursery/infant school time was devoted to free play/free choice activities. The junior group received 8·8 per cent and the senior group 2·7 per cent of the time devoted to free play/free choice.

9. The average number of minutes of actual lesson time per school week was 925. Nursery/infants received 41 per cent of this time in free play/free choice activities but a dramatic fall in time for these activities occurred with the older age group, i.e. 18 per cent for juniors and 5 per cent for seniors.

10. Seventy-eight (89%) headteachers indicated what they considered were their children's needs in order of priority. The largest number, 42 (54%), indicated social competence; 26 (33%) indicated school facilities, and 22 (28%) indicated language development.

B. *The skills of 151 mentally handicapped children*

1. The mean percentage scores assessed for the four main areas of the Skills-Assessment Sheet showed that *self-help* skills (74) were the easiest and *communication* skills (40) the most difficult. The other scores were 59 for *occupation* and 42 for *socialization*.

2. Many children lacked those skills necessary to function normally in the community. For example, the weakest skills in *self-help* were 'walks about neighbourhood and crosses streets (unsupervised)' (29%), and 'uses public transport independently' (7%), and in *socialization*—'goes on simple errands outside the home and school' (23%) and 'makes minor purchases' (21%).

3. The mean percentage successful responses in the skills followed a pattern one would expect with the increase in chronological age, maturity and intellectual functioning. The rise in percentage responses was sharp at first (up to the 11–14 age group) but then the rise fell away for the over 14 age group and there was no evidence to suggest that over 14 age group behaved differently from the 11–14 group—there was a levelling off in achievement.

C. *Information from 52 parents*

1. Sixty per cent of parents were informed of their child's handicap before one year had passed; 23 per cent between one and three years; and 17 per cent between 4 and 7 years. Only 8 parents (15%) received adequate advice and help from their doctors.

2. Twenty-eight parents (54%) were visited by a Mental Welfare Officer/Social Worker before school and after the child had started school. Eleven of these parents stated that they had received adequate advice and help. Forty parents (77%) were visited by a Health Visitor; 16 of these stated that they had received adequate advice and help. The majority of parents had to seek help.

3. Thirty-two (61%) children attended school before 5 years of age. Forty-nine (94%) travelled to school in special transport provided by local authorities. Thirty-two (62%) children were away from home 7 hours or more (12 of these were away 8 hours or more).

4. Seventy-three per cent of parents stated that their greatest concern was that their children required constant attention and supervision especially during the early years.

5. Thirty-three (63%) parents were satisfied with their children's schools. Twelve 'dissatisfied' parents of 'high grade' children suggested that their children should either be in a special school for 'higher grades' or in special classes in the normal school.

6. Sixty-three per cent of parents regarded 'speech development' as their child's greatest need in school; 25 per cent stated 'reading and writing' and 8 per cent 'social competence'.

7. Only 48 per cent of the schools had parent-teacher associations. Only 2 parents (4%) had visited a school less than two months ago; 31 (60%) three months ago; 9 (17%) 6 months ago; 8 (15%) 12 months or more ago, and 2 (4%) had never visited a school.

Discussion of the results

A. *The new day special schools*

The fairly high percentage (35%) of mongols found in the present investigation is probably a reflection of the incidence of mongolism throughout the British Isles (Tizard, 1967; Marshall, 1967). This is important because, generally speaking, non-mongols exhibit a wide range of ability and atypical behaviour whereas, generally speaking, most mongols share common characteristics and may be classified as low, medium and high. Furthermore, the fact that mongols are more easily recognised at birth is of particular importance because advice may be given to parents at an early age and so they have a longer period of adjustment to the child's handicap by the time that he goes to school.

The results of the investigation suggest an improvement in adult/pupil ratio when one compares these results with those of other investigations (Marshall, 1967; Norris, 1968), but the ratio is still poor.

There appears to have been a substantial improvement in the number of qualified staff when the present findings are compared with official figures (Ministry of Health 1966; DES 1970). This increase in qualified personnel has been due to several factors, the most important being the increase in training courses since the Scott Report was published in 1962, and the establishment of the Central Training Council for Teachers of the Mentally Handicapped in the same year.

The findings suggest a substantial increase in purpose-built schools during the last 25 years. Most schools had adequate playground facilities and many had, in addition, either playing fields or extensions to 'hard' playgrounds. But even though there are indications of improvements in facilities, many schools considered that they had inadequate storage facilities both indoors and out. There appears to be a need to study the effectiveness of the designs of certain special schools for the mentally handicapped.

There appeared to be an adequate supply of general classroom equipment but there were many schools without appropriate gymnastic equipment and an apparent inadequacy in facilities for domestic arts, handwork and gardening, especially for older children. This applied to many purpose-built schools as well as the others. There is an urgent need for advisers in the above subjects to spend some considerable time in the schools.

Only 58 per cent of schools had facilities for the 'under-fives' and this suggests that there is an urgent need for an increase in facilities not only in special schools but also in the normal nursery class/school and infant school for these children.

A surprisingly small percentage of schools (52%) had parent-teacher associations; and this did not appear to be compensated by the close, informal school-parent relations which are essential for the optimal development of the child.

There was a general lack of flexible free play/free choice type of curriculum in the sample schools. The majority of schools used formal timetables. Forty-nine per cent of the school week in 40 schools was allotted to non-teaching activities (i.e. administration, assembly, toilet, rest periods, meal times and breaks). This was only slightly lower than the information obtained by other investigators (Simpson, 1967; Tizard, 1967; Marshall, 1967; Norris, 1968) and may be still too high.

The rather drastic fall in free play/free choice activities for older children suggests that the older the child the more he is 'kept at his desk', involved in sedentary activities at the expense of motor activities and a 'learning by doing' approach.

Over half the headteachers regarded social competence as their children's greatest need, and this is reflected in the low scores for many skills involved in social competence already referred to. Obviously headteachers believed that they needed much more in the way of school facilities, and this need was placed above language development.

B. *The skills of 151 mentally handicapped children*

Even though the highest mean percentage scores assessed were in the area of self-help, only 29 per cent could walk about the neighbourhood and cross the streets unsupervised, and only 7 per cent could use public transport independently. This suggests that there are many mentally handicapped children who are unable to function in the community normally. This suggests 'over-protectiveness' on the part of the home and the school, even when one appreciates the traffic dangers in a modern urban community. The inability of many children to function in the community is further reflected in the area of socialisation, where it was found that many were unable to go on simple errands. The apparent need for more realistic situations at the class shop and at real shops is reflected in the finding that none of the youngest was able to make minor purchases, and only 14 per cent of the 8–11 group and 37 per cent of the remaining older children could do so.

None of the youngest children (under 8 and 8-11) could name coins up to 10p, and only 27 per cent of the remaining older children could do so. None of the youngest could give change out of 5p, and only 21 per cent of the older children could do so. These findings are very disquieting, especially in the case of older children. If mathematics is to be functional in that the child should be able to apply it within his environment, then proficiency in simple matters is urgently needed.

The mentally handicapped child's weakness in language development is indicated in the lowest mean percentage score for the four areas, i.e. communication.

There was no evidence to suggest that the older children (over 14 group) behaved differently from the children in the 11–14 group. The reasons for this may well be a culmination of factors such as a drastic fall in the percentage of time allotted to free play/free choice activities for older children with low mental ages; lack of facilities in domestic arts, handwork, gardening and PE, and less time devoted to manipulatory skills and muscular co-ordination; lack of opportunities to exercise socialisation skills; 'over-protectiveness'; a general lack of 'learning by doing', and lack of appropriate teaching methods. However, even though the above may well be contributing factors, the reason for the apparent lack of improvement in skills on the part of the older children may well be twofold in that teachers may be satisfied with what has been achieved by the age of 13 or 14, and do not demand much after that. This points to the need for higher levels of expectancy. But one must also consider the possibility of a rather long pause after a particular age before measurable progress can be recorded (Gunzburg, 1968).

C. *Interviews with the parents of 52 children*

There appears to be an improvement in informing parents as early as possible of their child's handicap when the present findings are compared with those of other investigations (Mittler, 1954), but there were still complaints from parents who suggested that they should have been informed of their child's 'obvious' handicap, (e.g. mongolism) at a much earlier time. Furthermore, the findings suggest an urgent need to provide parents with appropriate education so that they fully appreciate the implications of having a mentally handicapped child, although there was no real evidence that the majority of families suffered from intolerable problems.

The number of long, fatiguing journeys to school should be reduced by building more hostels near schools together with more educational provision in the normal primary school for certain mentally handicapped children.

In spite of frequent inadequacy on the part of social services, the majority of parents (77%) exhibited resilience and resourcefulness. However, there does appear to be a need for voluntary agencies to arrange baby sitters to help these parents.

A call for more integration rather than segregation came from many parents. It is suggested that normal primary schools should be prepared to take certain mentally handicapped children, and investigations should be carried out into the feasibility of more integration.

The majority of parents emphasized what has become recognized as one of the mentally handicapped child's greatest weaknesses, i.e. language. Only 8 per cent regarded social competence as their child's greatest need; this may be a reflection of the parents' 'over-protectiveness', or, on the other hand, a recognition that a better mastery of language would contribute to an increase in social competence as well as being desirable in itself.

References

Cain, L. F. and Levine, S. (1961) *A Study of the Effects of Community and Institutional School Classes for Trainable Mentally Retarded Children.* San Francisco State College.

Department of Education and Science (1971) *Reports on Education No. 69.* London: HMSO.

Doll, E. A. (1965) *The Vineland Social Maturity Scale.* Education Test Bureau, Educational Publication Inc. USA.

Gunzburg, H. C. (1963 and 1965) *Progress Assessment Charts.* London: National Association for the Mentally Handicapped.

Gunzburg, H. C. (1968) *Social Competence and Mental Handicap.* London: Baillierè, Tindall & Cassell.

Kirk, S. A. (1965) 'Research in Education'. In: Stevens, H. A. and Heber, R. (Eds). *Mental Retardation: A Review of Research.* Chicago: Chicago University Press.

Lunzer, E. A. (1965). *Manchester Scales of Social Adaptation.* Slough: NFER.

Marshall, A. (1967). *The Abilities and Attainments of Children Leaving Junior Training Centres.* London: National Association for the Mentally Handicapped.

Ministry of Health (1967) *Annual Report of the Ministry of Health for 1966.* London: HMSO.

Mittler, P. (1954) *The Mental Health Services* (Series no. 252). London: Fabian Society.

Neale, M. D. and Campbell, W. (1963) *The Education of the Intellectually Limited Child and Adolescent.* Melbourne: Novak.

Norris, D. (1968). *Some Observations on the School Life of Severely Retarded Children.* A Monograph Supplement to the *Journal of Mental Subnormality.*

Russell, R. W. (1964) *A Program of Special Classes for Children with Learning Disability.* New Jersey: Association for Brain-Injured Children.

Scott Report (1962) *Report of the Sub-Committee on the Training of Staff of Training Centres for the Mentally Subnormal.* London: HMSO.

Simpson, P. F. (1967) 'Training centres—a challenge', *Spec. Educ.*, 56, 34–8.

Tizard, J. (1967) *Survey and Experiment in Special Education.* London: Harrap.

Book Reviews

Running a School

by C. H. Barry and F. Tye

Maurice Temple Smith, London, 1972, 248 pp.,

£2·50

This book is based on the results of an investigation into existing practice. It also includes an infusion of managerial ideas from areas outside education. The book is divided into four sections: Principles, Organization, Resources and Assessment, each section being highly detailed and well ordered. It is a scholarly work and in spite of the authors' expressed hope that it will be bought and read by laymen as well as professionals it is nevertheless, in my view, strictly for professionals. It is not light reading; it is one of those books which needs to be re-read as the early chapters gain in relevance after the whole book has been studied. As may be imagined, the section on organization is the largest of the four sections and the topical themes of consultation, delegation and administration are well examined within it.

In many ways the contents of this book will appear more relevant to those actually running schools than to those preparing themselves for such management. This is because much of the text acts as a timely reminder of what the head-teacher should be doing. For example, considerable emphasis is placed on assessment of what is being done as a basic operation of good management.

The authors make it quite clear that their work is more concerned with general principles than with specific detail. Consequently this is not a handbook to refer to for suggested ways of carrying out specific tasks. In some ways this is a loss. However, very few educational writers today are prepared to nail their colours to the mast in this respect so we should not be surprised. The general principles are nevertheless widely embracing and sensibly examined, leaving very little room for debate.

The chapter on the assessment of staff is hard-hitting, but the arguments advanced are unassailable. The question of choice in decision making is powerfully argued and the outcomes and consequences that must be accepted are realistically drawn.

Chapter 6, misleadingly entitled 'A Personal View', concludes the book's first section on principles. It is a chapter which stands alone and summarizes the aims and objectives relevant to most secondary schools today. It is the chapter on which the rest of the book hinges and consequently should be read closely. In this context it is refreshing to find authors who do not persistently hedge their bets with frequent reservations. Barry and Tye state their contentions with conviction and having identified the basic principles which underlie the running of a secondary school boldly put them into priority order.

The chapter on 'Tackling the Job' is a short working guide to which every head-teacher could well refer every time a major decision has to be made. The cautionary, self-questioning approach to decision-making outlined in this chapter can only lead to better decision-making.

There are few reservations I would have about this book. It is a fine guide, dealing in a practical way with the fundamental educational principles that emerge in the running of a school; and it is certainly one of the more scholarly works of its kind that has come my way in recent years.

JOHN PRICE

Education or Domination

edited by Douglas Holly,

Hutchinson Publishing Group Ltd,

Paperback £1

Presumably all the contributors to this collection see themselves as educational radicals, and presumably too they all started dedicated to 'dispelling mists' and penetrating the verbiage of most educational writing—an exciting prospect for the reader, ever hopeful of finding the road to educational utopia, the radical thesis to back all others.

So, to the first chapter ('Education and Politics' by Ted Benton)—and alas, right back into a Force Five mist of Marxist analysis through which dimly shines the message that most radical directions in education are naive fiddling as long as they remain in the context of a capitalist society. Everyone in the Labour Party and to the left of it is naive or occasionally worse if they succeed in educating working class children out of thinking of society in terms of class divisions (I always thought the battle was going the other way). His precription for the future is predictably misty: some, strands include *Rank and File's* encouragement of school and student unions, trade union action in defence of victimized militants in education (Who is victimized? Who is a militant?), the use of university facilities for flying pickets during the miners' strike (guest places at the high table? A quick PhD in aeronautics?). And where do you suppose Mr Benton is? It would have been better if the editor had maintained the suspense by putting potted biographies at the end of contributions. Mr Benton teaches in the sociology department of Essex University (or he did when he wrote this essay).

Working in a free school, it was hard to resist the temptation to start this book from the back, with Guy Neave's essay on 'The Free Schoolers'. Benton's verbiage lowered my resistance, so I turned to the back, only to find my apathy turned to apoplexy. Mr Neave is a fervent supporter of The Comprehensive Ideal. He has even done some research to support his support. But all these nasty radicals have come along and said that comprehensive reorganization isn't necessarily radical at all. So he is on the warpath, determined to squash at least one non-conformist school. In his essay, Mr Neave manages to compound all the current misconceptions about free schools, and add a few new ones of his own. Of course, it is true that in any new and decentralized movement, there are bound to be conflicting strands and a lack of concensus about aims and methods, but if Mr Neave had come as far as Islington before setting pen to paper, he would have had to rethink some of his generalizations (and, hopefully, his whole educational philosophy).

The current development of urban free schools in Britain (similar to one strand of the free school movement in America) is not a drop-out philosophy, but a conscious effort to provide a new institutional model free as far as possible from the constraints that inhibit learning in most schools. They see their job as influencing the direction of state education, and they are only 'outside' state education for as long as central authorities equate their role as dispensers of public money with an enormous measure of detailed control over what happens inside 'their' institutions. We are where we are in Islington because we believe what is wrong with schools is not redeemable simply by tinkering with the structures we have (appealing and comfortable though that sort of redemption must seem to those engaged in it). Comprehensive reorganization does not 'solve' anything, except removing a particularly vicious form of selection at 11. It is not a prescription for any particular kind of schools.

Mr Neave and very many others like him are obsessed with this 'comprehensive ideal', with making the system we've got work. He lambasts the free schoolers, for, amongst other things, being unable to 'create schools capable of dealing with large numbers of children', and suggests that 'Many features of the free school can and should be adapted to the state school.'

Most of the contributors to this book would agree on some of the basic constraints that make schooling anything but synonomous with education. It is sad that the conscious effort to promote an institutional model free of these constraints, which is what the current free school movement is all about, should be given such inaccurate and inadequate coverage. Mr Neave's view of educational change, from the dazzling heights of the University of Edinburgh Centre for Educational Sociology, is a sadly narrow-minded one: the sparse reading list he suggests may go some way to explain it.

For the rest, there is a valuable critique by Harold Rosen of Bernstein's linguistic theories of restricted and elaborated codes, which have been used over the years in support of both G. H. Bantock and Brian Jackson . . . so there must be something wrong somewhere. Rosen offers no alternative theory, merely a plea for more knowledge about the relationships of language to class.

Douglas Holly's own essay 'The Invisible Ruling Class', is at least not shrouded in the mists he condemns: he gives a clear and convincing analysis of the division of power in a capitalist society, and the way in which progressive movements in education as in other areas of social policy are tolerated by the tiny ruling class, and often encouraged by the liberal and conservative elements in the middle class—so long as the pattern of power is not unduly threatened (for those who want to see the reaction to any real threat to the power complex, turn again to Neave on the free-schoolers).

Holly is full of hopeful phrases about 'genuine' democracy and 'real' types of learning, but does not begin to suggest the institutional forms most likely to foster them. If he knew, perhaps he would not be still working in a University Institute of Education.

PETER NEWELL

Obedience to Authority
by Stanley Milgram
Tavistock Publications Ltd, 1974, 244 pp.,
Hardback £2·50

This most interesting book describes experiments carried out by Professor Stanley Milgram during the early 1960s. It deserves to be read by a wide audience, and not just by those with special interests in Social Psychology, as the issues discussed have relevance to us all.

Milgram feels that obedience, because of its ubiquitousness, has been overlooked as a subject of inquiry, and yet stresses its important role in shaping human action. An act, he points out, which is carried out under command is, psychologically, of a profoundly different character from action that is spontaneous. When people come to view themselves as no longer responsible for their own actions, and as instruments for carrying out another's wishes, then all the essential features of obedience follow. It is on these essential features that Milgram concentrates in his experiments, and he uses several real-life situations like the Nazi activities against the Jews and the My Lai massacre to illustrate the process of obedience in the world outside the laboratory.

The main aim of his experimental investigations was to find out when and how people would defy authority in the face of a clear moral imperative. Considerable detail is given about the basic design of the experiments and the variations that were added later, and it is clearly and carefully written. Two people come to the psychology laboratory to take part in a study of memory and learning. One of the two is designated as teacher, the other as learner and the experimenter explains that the study is concerned with the effects of punishment on learning. The learner is then conducted into a room, seated in a chair, with his arms strapped to prevent excessive movement, and electrodes are fixed to his wrists. He is told that he is to learn a list of word pairs and that whenever he makes a mistake, he will receive electric shocks of increasing intensity. Actually, in all the experiments the 'learner' is an actor and he doesn't receive any shocks, although this fact is not known by the 'teacher' who believes that the shocks are real. The focus of the experiment is the teacher, who is a genuinely naive subject and he can administer a shock ranging from 'slight' (15 volts) to 'Danger—severe shock' (450 volts).

The point of the experiments was to see how far a person would proceed in concrete and measurable situations in which he is ordered to inflict increasing pain on a protesting victim. At what point would the subjects refuse to obey the experimenter who appeared as a legitimate authority to whom the subjects felt some commitment?

Milgram's findings showed that, despite the fact that many subjects experienced stress and despite the fact that many protested to the experimenter, a substantial proportion continued to the last shock on the generator. In the post-experimental interviews when subjects were asked why they had continued to give shocks, a typical reply was—'I wouldn't have done it by myself. I was just doing what I was told'.

It seems that the subjects, unable to defy the authority of the experimenter, attributed all responsibility to him and Milgram points out that the prhase 'just doing one's duty' was heard time and time again in the defence statements of those accused at Nuremberg.

The author's theoretical analysis of the process of obedience lacks some of the rigour and coherence of the description of the experimental work and this is a weakness of the book. Milgram speculates in his analysis about what it is that makes a person who is usually decent and courteous act with severity against another person within the experiment. And his answer is that a person does so 'because conscience, which regulates impulsive aggressive action, is perforce diminished at the point of entering a hierarchical structure'. This hierarchical structure suppresses individual direction and control in favour of control from higher level components. Milgram doesn't discuss fully how his hierarchical structure is formed and operates by societal pressures. The world in which his experiments are conducted is that of the laboratory and yet he wants to make his findings more generalizable. This is a little difficult because written in the preface is:

'The laboratory problem is vivid, intense and real. It is not something apart from life, but carried to an extreme and very logical conclusion certain trends inherent in the ordinary functioning of the social world.'

From that statement, we assume that this extrapolation from the particular to the general is possible, so that the process of obedience as it occurs in the laboratory can be analysed similarly in other situations. However, later we read that—'If the experiment were to be carried on outside the laboratory, obedience would drop sharply', and this statement seems to imply that there is some difference, which is not clearly disscused, between experimental laboratory situations and real-life situations outside the laboratory. But such slight flaws do not detract from the importance of Milgram's work which should be read and discussed by anyone interested in the field of social interaction. The insight which he provides in a field of inquiry which has until now been almost ignored is very perceptive and thought-provoking. We shall all view some of our own actions and those of others differently after reading his book.

BARBARA ROBB

Playing, Working, Growing

by Geoffrey Ivimey

Maurice Temple Smith, 1974, 192 pp.,
Hardback £2·50

This book is written with the express aim of encouraging pre-school teachers to play a more active part in the development of their pupils than has been the accepted practice in the traditional English nursery school. This is not to say that Geoffrey Ivimey, a Lecturer at the London University Institute of Education, denies the importance of play, but he points out that whereas in traditional societies play was 'a recreation from the serious work of growing and living in the community', it has now become 'central to life' because children particularly in urban areas are cut off from the real everyday life of the community. What is necessary is for childhood to become once again, an apprenticeship for full participation in the life of the community, and the teacher or playgroup leader has an important role to play in bringing this about.

The author's subsequent discussion of how children learn appears to me to be oversimplified and somewhat contradictory. Having stated that children appear like animals to have a natural curiosity and desire to learn about their environment, he then goes on to state that motivation is not natural but has to be acquired. The ways the child acquires this are by reinforcement and imitation. Without denying the importance of these two processes, few psychologists today would agree that they are sufficient to account for all learning. However, the practical advice given on the experiences children need is very helpful. This is particularly so for the development of mathematical thinking where the author provides suggestions for sequences of activities which will help children develop the basic concepts. His discussion on the question of reading readiness is far more controversial. There is little or no evidence that a high proportion of 16-year-old illiterates 'have . . . been made to fail by being taught to read too early'. It is just as likely that reading never seemed important enough to them to persevere with it. However, the advice as to the kinds of pre-reading activities to provide for young children appears sound. (One might state in passing that a somewhat higher level of reading accuracy amongst the proof-readers would have been appreciated: the number of printing errors in the book being extremely irritating.) A final chapter on setting up a playgroup again shows the author's firm grasp of practicalities.

To sum up, the philosophy underlying this book is one which is basic to the improvement of pre-school education and many useful suggestions are offered to enable teachers to put this into practice. The theory has, however, frequently been oversimplified so as to reach a wider audience and at times dogmatic statements are made about areas where there is insufficient evidence to support them.

BARBARA L. THOMPSON

Shaping the Education of Slow Learners

by W. K. Brennan

Special Needs in Education Series, Routledge & Kegan Paul, 1974, 114 pp., Hardback £2·50,
Paperback £1·25

This book is the first in a series edited by R. Gulliford on children with 'special educational needs'. Brennan writes from an extensive background in teaching, research and curriculum development about the largest group of children considered by many professionals as requiring some form of special educational provision. In particular, Brennan has written for those following courses in special education and for administrators with responsibilities in this field. His purposes are to inform and to stimulate.

The book consists of five major sections. Initially the question 'Who are the slow learners?' is discussed. Various estimates of the incidence of such children in the population lead to the conclusion that the vast majority are currently attending ordinary rather than special schools. Some reservations are expressed as to whether the responsibilities that this implies are accepted by the majority of teachers. The low status at present accorded teachers of slow learning pupils is discussed.

Following this, Brennan reviews ideas concerning the characteristics of slow learners of most relevance to the teacher. Consideration is given to the implications of changing emphasis for those involved in developing curricula for the slow learner in both ordinary and special schools.

Brennan next presents a reformulation of the slow learners' distinguishing characteristics and hence of their educational needs. These are seen as met by three distinctive yet related aspects of education, namely, 'adaptive-developmental', 'corrective' and 'remedial'.

In the fourth section the progress of the slow learner on leaving school is discussed largely on the basis of American and British research. Failure in post-school adjustment is seen as mainly attributable to the slow learners' emotional immaturity and impulsivity coupled with the failure of fellow-workers and employers to provide the support that a proportion of these adolescents require. This leads to a discussion of the curriculum required if the slow learners' chances of adequate post-school adjustment are to be optimized.

Finally Brennan discusses the curriculum of the slow learner in terms of the relationship between skills which the pupil needs to master and those broader aspects of education that provide the essential context within which such competencies can be developed. He concludes with a plea for the involvement of the pupil in the process of curriculum development.

A comprehensive bibliography is provided. It will enable the interested reader with access to a good reference library to follow up in depth the many points of controversy, the extensive practical suggestions and the research findings discussed in the text.

Having worked for a number of years with slow learning children, the reviewer is appreciative of the considerable contribution Brennan has made, and is making, in this field. Despite this, some disquiet remains about certain of his arguments. In part this is because of the somewhat arbitrary nature of the definition of the slow learner. It can be argued that we have principles of education applicable to all children and that, at this level, all children have similar educational needs although individual differences between pupils determine how best these needs might be met. Thus one finds that whenever Brennan suggests what appears to be in the best interests of slow learning pupils, these suggestions are almost inevitably equally valid for all pupils. For example, one could question whether, in helping slow learning pupils to overcome their learning difficulties, '... teachers individualize their methods to a point which would be neither possible nor desirable with normal pupils ...' (p. 67). Extensive use is made of the term 'needs' and the importance of their definition is stressed but insufficiently dealt with other than at a pragmatic level.

The text is replete with pertinent references. Some readers unfamiliar with the style might find the writing rather 'dense' in this respect. Students undertaking further study in this area will find that these references provide a key to the detailed expansion of both theoretical and practical aspects of the education of slow learning children considered by the author.

The book is not light reading. Whilst it can provide a valuable background to the tyro, it is probably more readily comprehended after some experience of work with slow learning pupils. In its 114 pages Brennan has packed a tremendous amount of recent historical information, research findings, argument and constructive suggestion concerning the education of slow learning children. Such a book will contribute to improving the morale and professional self-esteem of those teachers of slow learning pupils who read it. For those who are not concerned with the daily work of educating slow learning pupils, it will help their appreciation of the professionalism that can be found.

The decision, presumably editorial, not to include a subject and an author index is to be regretted. It is hoped that subsequent publications in this series will contain both.

P. D. Pumfrey

Teaching Languages to Adults for Special Purposes, CILT Reports and Papers II

Centre for Information on Language Teaching and Research

1974, 100 pp., Paperback £1·50

This report of papers presented at an invited conference; London, 5–7 July, 1974 includes six main papers; brief accounts of eight actual courses and a review of recent surveys of the foreign language needs of adults in Britain, Sweden and Germany.

The need for thorough research and the impossible demands made by outside institutions unable to accept that learning foreign languages is difficult are illustrated by Mabel Sculthorp and B. Fitzjohn. The development and evaluation of course programmes are discussed by Marcel Kuhn and George Perren respectively, while in the most interesting paper J. L. M. Trim describes a model for the analysis of language—using situations evolved as part of the Council of Europe adult language learning project.

The accounts of courses include Chinese and Japanese as well as Russian German and French. They are interesting in that they tell us a little of what various institutions are doing, but are all frustrating in the total lack of examples of actual teaching material. Descriptions of methodology never seem to include details of what is actually on the pieces of paper and lengths of tape used. Without such evidence it is impossible to judge the true interest of new material.

The appendices listing current research and a select bibliography are most useful.

D. R. SCARBROUGH

The Education of the Poor: the History of a National School 1824–1974

by Pamela and Harold Silver

Routledge & Kegan Paul, 1974, 197 pp. Hardback £3·95

In his preface to this case history of a London primary school, Brian Simon notes that, whereas there have been upward of 500 histories of public and endowed grammar schools, scarcely one has been published of the 19th century elementary schools in which most of the nation's children were educated. For information about the latter we have leaned heavily up till now on the no doubt often caricatured and overcolourful descriptions of the Victorian novelists. So Pamela and Harold Silver have performed a valuable service in making an objective contribution to the correction of the imbalance. They were assisted in their study of St. Mark's School, Kennington, by two atypical features—its occupation of the same premises throughout the 150 years of its life and the availability of extensive records reaching back beyond the introduction of log books under the Revised Code to its foundation as a National School in 1824.

The authors are at pains to relate the school's development to local and national social changes and, more specifically, to the evolution of a statutory system of primary and secondary education for all. Of particular interest is their outline of the similarities and differences in the definitions of aim and purpose offered by the school's policy makers over the years and the effect that changes in definition had on the curriculum. The National Society for Promoting the Education of the Poor in the Principles of the Established Church, founded in 1811 to meet the needs created by industrialization and urbanization and the challenge of the new radicalism, was quite unambiguous about its intentions in regard to the million children who were attending its 17,000 day schools by mid-century. Its aim was 'to communicate to the poor . . . such knowledge and habits, as are sufficient to guide them through life, in their proper stations, especially to teach the doctrine of Religion'. Religion continues to bulk largely in the pronouncements of the 1970s. St. Mark's remains a church school, and its current prospectus for parents lays marked stress on religious and moral education; but, the authors observe, it is no longer 'by definition' for the 'instruction of the infant poor', although most of its pupils are still working class; and its stated aim is to help them fulfil their individual potentialities rather than to confine them to the 'knowledge and habits' considered appropriate to their 'proper stations'. This shift in purpose was already manifest by the end of the 1860s when a small number of public and grammar schools began to make scholarships available for elementary school children, and lower middle class parents were, in consequence, prepared to send their children to the better elementary schools like St. Mark's. By this time, also, the 'ragged schools' had taken over the task of providing London's most destitute children with a minimal education.

'Useful and religious knowledge' 150 years ago meant essentially scripture and reading, writing and arithmetic, with needlework, in addition, for the girls, and it was purveyed by Lancaster and Bell's elaborate monitorial system of rote learning and rewards and punishments. This stark regime was slow to alter, but with the shift in aim came diversification of the curriculum. Geography, history and grammar were being taught by the 1840s, and in 1911 prizes were awarded for, in addition to the subjects already mentioned, spelling, recitation, drawing, and science. As for methods, payment by results introduced by the Revised Code ended in the 1890s, and the 1920s witnessed the triumph, in theory at least, of individualized learning over routinized instruction.

Given the dearth of case histories such as the one the Silvers have written, it is unfortunate that theirs should also disappoint. Despite liberal, and often graphic, quotations from the school log book and other contemporary documents, its coverage of 150 years in 170 pages is somewhat sketchy and pallid. Crucially it fails to convey the real feeling of what being at school was like for most children in the last century. For this we still have to turn to the fictional and autobiographical accounts of the major, and indeed minor, Victorian novelists.

ROBERT JEFFCOATE

Information

Educational Research is published in November, February and June for the National Foundation for Educational Research by the NFER Publishing Company Ltd. (Book Division, 2 Jennings Buildings, Thames Avenue, Windsor, Berks. SL4 1QS).

Annual Subscriptions. UK: £4·00 post free. Overseas (excluding USA and Canada*): £5·00 post free.

Subscriptions date from issue number 1 of each new volume (November). New subscribers in the period November to June will receive all three issues of the *current* volume, unless they specifically state that their subscription is for the *next* full volume.

Single Copy Prices. UK: £1·50 post free. Overseas*: £2·00 post free.

Payment should accompany all orders except where separate arrangements have been made. Payment from overseas should be in sterling.

Orders should be sent to NFER Publishing Company Ltd. (Journals Department), at the address above.

*Sole distributors for the USA and Canada: APS Publications Inc., 150 Fifth Avenue, New York, NY 10011, USA.

Advertisement Inquiries should be made to the NFER Publishing Company Ltd. (Book Division).

Editorial Correspondence. Manuscripts for consideration, and books for review should be sent to The Editor, Educational Research, National Foundation for Educational Research, The Mere, Upton Park, Slough, Berks. SL1 2DQ.

Notes for Contributors. Authors are invited to submit reports of original educational research, reviews of recent research in particular areas, or discussion articles on research topics. Articles should begin with a brief summary, and should not normally exceed 3,000 words. As well as the intrinsic interest of the article, conciseness and clarity are important considerations. Technical jargon should be avoided, and, where possible, statistical data should be summarized in the text, although tables may be included if they are clearly presented. Authors are encouraged to describe their findings in terms intelligible to the non-expert reader.

Articles should be typed in double spacing on one side of the paper, and should be submitted in duplicate. Drawings should be ready for direct reproduction. References should be in the style used in this issue. A manual of advice to authors is available from the Editor.

Only material which has not been published elsewhere can normally be accepted.

Publication rights are retained by the Editor. Reproduction in whole or in part is forbidden except:

1. Under the terms of the Fair Copying Declaration of the Royal Society, or
2. When written permission has been given by the Editor.

The conclusions drawn and the opinions expressed in papers are those of the authors. They should not be taken to represent the views of the NFER.

© National Foundation for Educational Research in England and Wales, 1975

Printed by King, Thorne & Stace Ltd., Kingthorne House, School Road, Hove, Sussex BN3 5JE

L. G

25 SEP 1975

In this issue

Further education: an alternative route reviewed
Standards of reading: a critical review
Measurement of human values: the work of Milton Rokeach